



PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS
Programa de Pós-graduação em Engenharia Elétrica

Rosemary Antonia Lopes Faraco

Detecção de Falhas Elétricas em Motores de Indução Utilizando Redes Neurais

Dissertação submetida à banca examinadora, designada pelo Colegiado do Curso de Pós-Graduação em Engenharia Elétrica da Pontifícia Universidade Católica de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Ciência.

Orientador: Professor Dr. Pyramo P. da Costa Júnior.

**Pontifícia Universidade Católica de Minas Gerais
Belo Horizonte
2000**

Aos meus pais, Alberico e Lúcia (*in memoriam*).
Ao meu querido esposo Renê e ao amado filho Felipe.

AGRADECIMENTOS

A Deus pela força e coragem para vencer todas as dificuldades.

Ao Professor Dr. Pyramo P. da Costa Júnior, orientador.

Aos Professores Frederico R. B. da Cruz e José Celso pelo apoio e valiosas sugestões.

Ao Prof. e coordenador do curso Dr. Luiz Danilo Barbosa Terra, pelo apoio e incentivo.

Ao Prof. Dr. Roberto, pelo sorriso e brincadeiras que suavizaram a luta.

Ao Sr. Alci Barbassa e Sra Lúcia de Aguiar, gerentes de RH da Copasa, pela amizade, estímulo, apoio e, acima de tudo, pelo voto de confiança.

Ao Prof. Alexandre, chefe do Depto de Engenharia Elétrica e Automação pelo apoio e atenção dispensados.

Ao Prof. Bustamante, Diretor do IPUC, pelo apoio e amizade.

Ao Felipe, meu filho querido, motivador de meus maiores esforços.

Aos ex-bolsistas José Carlos, Adriano Alves e Bruno César pelos trabalhos realizados durante as respectivas bolsas de Iniciação Científica.

Aos demais Professores, funcionários e aos colegas do Mestrado, pela solidariedade e colaboração a mim dedicada em momentos difíceis.

À FAPEMIG, pelo fomento.

“Para triunfar es necesario vencer, para vencer es necesario luchar, para luchar es necesario estar preparado, para estar preparado es necesario proveerse de una gran entereza de ánimo y una paciencia a toda prueba. Esto requiere, a su vez, llevar constantemente a lo íntimo de la vida el incentivo de la suprema esperanza de alcanzar aquello que se anhela como culminación feliz de la existencia.”

Carlos Bernardo Gonzalez Pecotche

RESUMO

Estudou-se e desenvolveu-se um sistema para identificar faltas elétricas nos enrolamentos do estator de máquina de indução trifásica, utilizando redes neurais artificiais. Inicialmente realizou-se uma revisão histórica de métodos quantitativos e qualitativos estudados e desenvolvidos para detectar e localizar faltas em máquinas elétricas a partir das medidas de variáveis elétricas, mecânicas, de temperatura e análise química. Em seguida, foi feita uma revisão teórica sobre redes neurais artificiais, tendo-se em vista a sua aplicação no reconhecimento de padrões, identificando faltas em máquinas de indução. Dentro desta revisão estudaram-se quatro diferentes arquiteturas de redes neurais artificiais, que foram implementadas utilizando-se o “Toolbox” de Redes Neurais do MATLAB 6.5. Tais redes foram treinadas e testadas, utilizando-se uma base de dados, obtida a partir das medições de variáveis elétricas num protótipo de máquina de indução. O protótipo forneceu o ambiente para simulação em laboratório, das condições de funcionamento normal e sob faltas nos enrolamentos das fases do estator da máquina de indução. A partir dos resultados obtidos no treinamento e teste das diferentes redes neurais, analisou-se o desempenho de cada arquitetura estudada, identificando-se a mais eficiente para se construir um sistema capaz de identificar e localizar faltas no enrolamento do estator de uma máquina de indução, bem como condições normais de funcionamento.

ABSTRACT

It was studied and developed a system to identify electric faults in the turns of the winding from a three-phase induction machine, using artificial neural networks. Initially was made a historical revision of published quantitative and qualitative methods for to perform the faults detection in electric machines based on direct measurements of electric variables, mechanics variables, temperature and chemical analysis. After that, it was made a theoretical revision about artificial neural networks, with the aim of to know patterns recognition methods and their application in the identification of faults in the induction machine. Inside of this revision were studied four different architectures of artificial neural networks, which were implemented using the neural networks toolbox of MATLAB 6.5. The measurements of electric variables in a prototype of a three-phase induction machine, under different faults and not faults conditions produced the necessary data to train and test the neural networks. Analyzing the results obtained in the training and test of the different neural networks, was identified the neural network that yielded satisfactory performance to build an electric faults detection system. This is capable to identify the situations for which it has been trained and signaling with the identification of the defect as well as of normal conditions of operation from a three-phase induction machine.

LISTA DE ILUSTRAÇÕES

FIGURA 1 – Estrutura generalizada dos métodos de detecção de faltas baseados em modelos de processos, grandezas mensuráveis e não mensuráveis.	39
FIGURA 2 – Motor de indução trifásico, com rotor bobinado, utilizado como simulador físico de faltas, neste trabalho.	80
FIGURA 3 – Adaptações feitas no motor mostrado na FIGURA 2, para disponibilizar externamente os terminais de cada uma das bobinas que compõem o enrolamento.	82
FIGURA 4 – Detalhe da Figura 3, mostrando terminais externos das bobinas 7-8 até 15-16 do estator e os terminais rotóricos.	83
FIGURA 5 – Esquema de ligação para simulação de falta em uma bobina da fase “A”.	87
FIGURA 6 – Esquema de ligação para simulação de falta em duas bobinas, em duas fases diferentes, porém, ao mesmo tempo.	87
FIGURA 7 – Esquema de ligação do estator para simular falta em três bobinas, uma por fase, ao mesmo tempo.	88
FIGURA 8 – Curva obtida a partir da interpolação da curva tensão x corrente a vazio.	95
FIGURA 9 – Treinamento com parada antecipada.	100
FIGURA 10 – Visualização de tela de saída de programa feito em MATLAB, demonstrando, através da curva senoidal, como uma rede faz a generalização, utilizando a Perceptron Multicamada.	101
FIGURA 11 – Diagrama do processo de treinamento da rede na Simulação 1 (Rede <i>Backpropagation</i>).	108
FIGURA 12 – Evolução do treinamento da rede na Simulação 1 (Rede <i>Backpropagation</i>).	109
FIGURA 13 – Trecho transcrito da tela de simulação 1, apresentando a metodologia de minimização de erros, o número de épocas, o erro médio quadrático, a norma do gradiente e a finalização da simulação.	110
FIGURA 14 – Histograma de erro absoluto na fase A.	112
FIGURA 15 – Histograma de erro absoluto na fase B.	112
FIGURA 16 – Histograma de erro absoluto na fase C.	113
FIGURA 17 – Histograma do erro absoluto considerando a somatória de erros nas 3 fases.	113
FIGURA 18 – Diagrama do processo de treinamento da rede na Simulação 2.	118
FIGURA 19 – Evolução do treinamento da rede na simulação 2.	119
FIGURA 20 – Trecho transcrito da tela de simulação 2, apresentando a metodologia de minimização de erros, o número de épocas, o erro médio quadrático, a norma do gradiente e a finalização da simulação.	119
FIGURA 21 – Histograma de erro global de desempenho - Simulação 2.	120
FIGURA 22 – Diagrama do processo de treinamento da rede na Simulação 3 com a Rede <i>Backpropagation</i>	125
FIGURA 23 – Evolução do treinamento da rede na simulação 3.	126
FIGURA 24 – Trecho transcrito da tela de simulação 2, apresentando a metodologia de minimização de erros, o número de épocas, o erro médio quadrático, a norma do gradiente e a finalização da simulação.	126
FIGURA 25 – Histograma de erro global.	128
FIGURA 26 – Diagrama do processo de treinamento da BAM.	133
FIGURA 27 – Histograma de Erro no Treinamento da BAM.	134
FIGURA 28 – Histograma de Erro na Validação.	135

FIGURA 29 – Processo de seleção dos atratores e avaliação dos padrões para a rede Hopfield.	139
FIGURA 30 – Diagrama do processo de treinamento da Hopfield	140
FIGURA 31 – Histograma de Erro na Validação	142
FIGURA 32 – Diagrama do processo de treinamento da rede GRNN	145
FIGURA 33 – Histograma de erro global, apresentando a quantidade de BITS incorretamente classificados.	147
FIGURA 34 – Formato do arquivo <i>batch</i> de dados.	163
FIGURA 35 – Tela 1 do sistema desenvolvido em MATLAB	164
FIGURA 36 – Tela apresentando as opções de redes <i>feedforward</i> .	164
FIGURA 37 – Tela disponível para cada opção de rede, apresentando as operações disponíveis.	165
FIGURA 38 – Tabela contendo índices de desempenho do treinamento da rede	165
FIGURA 39 – Unidade Neural Básica de McCulloch e Pitts	171
FIGURA 40 – Funções de Ativação Típicas	172
FIGURA 41 – Estrutura de uma Rede Neural Artificial Multicamadas	176
FIGURA 42 – Ilustração da estrutura da Memória associativa bidirecional – BAM, similar à rede Hopfield.	192
FIGURA 43 – Estrutura da Rede Neural Hopfield	194
FIGURA 44 – Arquitetura da Rede Neural de Base Radial – RBF.	201
FIGURA 45 – Exemplo de uma função Gaussiana gerada a partir da equação (78).	203
FIGURA 46 – Arquitetura da Rede Neural de Base Radial – RBF.	204
FIGURA 47 – Arquitetura da Rede Neural de Base Radial tipo GRNN.	209
FIGURA 48 – Arquitetura da Rede Neural de Base Radial – RBF, tipo Rede Neural Probabilística.	210

INDICE DE TABELAS

TABELA 1 – Síntese de metodologia relacionando as várias causas de faltas aos padrões de queima dos enrolamentos.	58
TABELA 2 – Ficha técnica do motor utilizado para obtenção dos dados experimentais	80
TABELA 3 – Conformação das bobinas do estator do motor de indução utilizado na pesquisa.	81
TABELA 4 – Dados eletromecânicos (elétricos e geométricos) do motor utilizado para obtenção dos dados experimentais desta dissertação	81
TABELA 5 – Dados eletromecânicos (geométricos) do motor utilizado como protótipo nesta dissertação.	82
TABELA 6 – Dados medidos durante o ensaio de rotor bloqueado.	84
TABELA 7 – Dados obtidos a partir do ensaio a vazio, com rotor aberto, frequência constante e tensão variável.	84
TABELA 8 – Ensaio de curto-circuito realizados no motor protótipo	85
TABELA 9 – Leiaute básico da base de dados gerada a partir dos ensaios realizados no motor protótipo.	89
TABELA 10 – Diferentes redes neurais estudadas em suas aplicações ao problema de faltas em motores de indução trifásicos.	92
TABELA 11 – Código que demonstra a utilização do recurso do MATLAB 6.5, denominado parada antecipada	98
TABELA 12 – Síntese das simulações realizadas.....	104
TABELA 13 – Leiaute da base de dados expandida.....	106
TABELA 14 – Exemplo de como foram somadas as ocorrências de erro absoluto	111
TABELA 15 – Distribuição do erro de validação para a simulação 1 com a Rede <i>Backpropagation</i>	111
TABELA 16 – Resultados da Simulação 1 com a Rede <i>Backpropagation</i>	114
TABELA 17 – Leiaute da base de dados expandida.....	116
TABELA 18 – Resultados da simulação 2 com a Rede <i>Backpropagation</i>	121
TABELA 19 – Leiaute da base de dados expandida.....	123
TABELA 20 – Desempenho Estatístico da Simulação 3 da Rede <i>Backpropagation</i>	127
TABELA 21 – Resultados da simulação 3 com a Rede <i>Backpropagation</i>	128
TABELA 22 – Resultados da Rede BAM.....	136
TABELA 23 – Resultados da Rede Hopfield.....	142
TABELA 24 – Desempenho Estatístico da GRNN durante a validação.....	146
TABELA 25 – Resultados da Rede GRNN	148
TABELA 26 – Resultados Comparativos	150
TABELA 27 – Medições feitas durante o ensaio a vazio.....	166
TABELA 28 – Medições feitas durante o ensaio em curto-circuito	167

SUMÁRIO

CAPÍTULO 1 – O PROBLEMA.....	15
1.1 – Introdução.....	15
1.2 – Formulação da Situação-Problema.....	17
1.3 – Objetivo do Estudo	18
1.4 – Justificativas	19
1.5 – Organização do texto.....	19
CAPÍTULO 2 – REVISÃO DA LITERATURA.....	21
2.1 – Introdução.....	21
2.2 – A detecção das faltas no contexto da manutenção	22
2.3 – Métodos para detecção e diagnóstico de faltas em processos produtivos	29
2.3.1 – O uso de modelos	30
2.3.2 – Métodos que utilizam reconhecimento de padrões.....	40
2.3.3 – Sistemas Especialistas	43
2.4 – Os defeitos em máquinas elétricas e suas causas	48
2.4.1 – ESTRESSE DE ORIGEM TÉRMICA.....	52
2.4.2 – ESTRESSE DE ORIGEM ELÉTRICA.....	53
2.4.3 – ESTRESSE DE ORIGEM MECÂNICA.....	55
2.4.4 – ESTRESSE DE ORIGEM AMBIENTAL (contaminação)	57
2.5 – Técnicas utilizadas na detecção de faltas em máquinas elétricas, baseadas em monitoramento de grandezas elétricas, magnéticas e mecânicas	60
2.5.1 – A detecção de faltas, utilizando técnicas de inteligência computacional e o método qualitativo	71
2.5.2 – A utilização de Redes Neurais no monitoramento das condições das máquinas elétricas.....	73
2.6 – Conclusões.....	77
CAPÍTULO 3 – DESENVOLVIMENTO DO PROTÓTIPO E MONTAGEM DA BASE DE DADOS	78
3.1 – Introdução.....	78
3.2 – O simulador físico de faltas.....	78

3.3 – Tratamento Experimental	83
3.3.1 – Simulação de curtos-circuitos nos enrolamentos do estator	85

CAPÍTULO 4 – MÉTODO QUALITATIVO DE DETECÇÃO DE FALTAS EM MOTORES DE INDUÇÃO, UTILIZANDO REDES NEURAIIS

4.1 – Introdução	90
4.2 – Considerações Gerais	91
4.3 – Escolhendo os tipos de redes neurais	91
4.4 – Pré-processamento da massa de dados	94
4.5 – Obtenção dos conjuntos de treinamento e validação	95
4.6 – Procedimentos gerais para melhoria das simulações das redes neurais artificiais	96
4.7 – Capacidade de Generalização	100
4.8 – Parâmetros para avaliação das redes	102
4.9 – Simulações e resultados obtidos	103
4.10 – Perceptron multicamadas com algoritmo <i>backpropagation</i>	104
4.10.1 – Simulação 1	104
4.10.1.1 – Estrutura da Rede <i>Backpropagation</i>	106
4.10.1.2 – Implementação feita	108
4.10.1.3 – Resultados do Treinamento	109
4.10.1.4 – Validação	110
4.10.1.5 – Conclusões da Simulação 1 com a Rede <i>Backpropagation</i>	114
4.10.2 – Simulação 2 (Rede <i>Backpropagation</i>)	115
4.10.2.1 – Estrutura da Rede	116
4.10.2.2 – Implementação feita	118
4.10.2.3 – Resultados do Treinamento	118
4.10.2.4 – Validação	120
4.10.2.5 – Conclusões da Simulação 2 com a Rede <i>Backpropagation</i>	121
4.10.3 – Simulação 3 com a Rede <i>Backpropagation</i>	122
4.10.3.1 – Estrutura (Simulação 3 com a Rede <i>Backpropagation</i>):	124
4.10.3.2 – Implementação feita	125
4.10.3.3 – Resultados do Treinamento	125
4.10.3.4 – Validação	127
4.10.3.5 – Conclusões da Simulação 3 da Rede <i>Backpropagation</i>	129
4.11 – Memória Associativa Bidirecional (BAM)	130

4.11.1 – Simulação 4 (BAM)	130
4.11.1.1 – Estrutura da Rede (BAM).....	131
4.11.1.2 – Implementação feita	133
4.11.1.3 – Resultados do Treinamento da BAM	134
4.11.1.4 – Validação da BAM	135
4.11.1.5 – Conclusões da Simulação 4	136
4.12 – Rede Hopfield.....	137
4.12.1 – Simulação 5	137
4.12.1.1 – Estrutura da Rede	137
4.12.1.2 – Implementação feita	140
4.12.1.3 – Resultados do Treinamento	141
4.12.1.4 – Validação	141
4.12.1.5 – Conclusões da Tentativa - rede Hopfield	143
4.13 – Rede de Regressão Generalizada (GRNN).....	143
4.13.1 – Simulação 6	143
4.13.1.1 – Estrutura da Rede	144
4.13.1.2 – Implementação feita	145
4.13.1.3 – Resultados do Treinamento	145
4.13.1.4 – Validação	146
4.13.1.5 – Conclusões da Tentativa.....	148
CAPÍTULO 5 – CONCLUSÕES FINAIS E PROPOSTAS DE CONTINUIDADE	149
5.1 – Introdução.....	149
5.2 – Resultados Comparativos de Desempenho das Redes Estudadas	150
5.3 – Conclusões Finais	151
5.4 – Metodologia para utilização prática da rede neural.....	152
5.5 – Propostas de Continuidade	153
REFERÊNCIAS BIBLIOGRÁFICAS	155
APÊNDICE A – SISTEMA DESENVOLVIDO	163
APÊNDICE B – DADOS OBTIDOS A PARTIR DOS ENSAIOS REALIZADOS NO MOTOR DE INDUÇÃO TRIFÁSICO	166

B.1 – Ensaio a vazio	166
B.2 – Ensaio de Curto-Circuito:	167
B.3 – Memória de cálculo dos parâmetros do motor de 4,6 kW	168
APÊNDICE C – REDES NEURAIAS ARTIFICIAIS	170
C.1 – Introdução	170
C.2 – Conceitos gerais	170
C.3 – Aprendizagem de uma Rede Neural Artificial.....	175
C.4 – Arquitetura: Perceptron multicamada	178
C.4.1 – Algoritmo de treinamento <i>backpropagation</i>	179
C.4.1.1 – Técnicas do <i>momentum</i> e medida da entropia do erro.....	183
C.4.1.2 – Variação da taxa de aprendizado	185
C.4.1.3 – O método <i>resilient backpropagation</i>	186
C.4.1.4 – O método do gradiente conjugado	187
C.4.1.5 – Métodos quasi-Newton	188
C.4.1.6 - Método de Levenberg-Marquardt.....	190
C.5 – Arquitetura: Memória Associativa Bidirecional (BAM).....	191
C.6 – Arquitetura: Rede Hopfield	194
C.6.1 – Considerações de Desempenho	198
C.7 – Arquitetura: Rede de Base Radial (RBF)	200
C.7.1 – Topologia característica	203
C.7.2 – Treinamento da rede RBF	206
C.7.3 – A Rede Neural de Regressão Generalizada - GRNN.....	209
C.7.4 – Topologia das Redes Neurais PNN - <i>Probabilistic Neural Network Architecture</i>	210

CAPÍTULO 1

O PROBLEMA

1.1- Introdução

As máquinas elétricas rotativas conquistaram um papel extremamente importante na vida moderna, não só nas indústrias como também dentro das residências. Assumiram tamanhos e potências diversos e podem ser encontradas em funções independentes ou como parte fundamental de processos altamente complexos. O que se espera é que tais máquinas tenham alta confiabilidade, vida útil longa e o mínimo de tempo reservado às manutenções. Com o objetivo de garantir estes quesitos é que os princípios de proteção foram desenvolvidos, no entanto, a proteção visa identificar e anunciar faltas que já estejam ocorrendo, não tendo como finalidade diagnosticar processos que conduzam a falhas ou minimizar a ocorrência de tais processos. Ou seja, destina-se a atuar no caso de defeitos, não a prevenir a ocorrência deles. A proteção pode atuar desconectando a máquina da sua alimentação. Esta ação é aceitável se a máquina elétrica estiver em funcionamento isolado ou não essencial a um processo produtivo. Se, contudo, este não for o caso, tal desligamento pode ser catastrófico para a produção.

Até a década de 60, não havia preocupação com programas de manutenção no setor industrial, a filosofia era a de parar quando quebrar, acarretando custos elevados que incluíam

perda de produção. Hoje, o panorama mudou nos setores produtivos. Com a exigência de maior competitividade, a operação estável dos equipamentos de um processo industrial tornou-se uma importante meta de produção e, como tal, passou a ser alvo de estudos. Assim o conceito de manutenção preditiva surgiu trazendo a filosofia de manutenção baseada no tempo, que usava dados estatísticos para determinar a troca incondicional de peças que poderiam ou não estar em boas condições de funcionamento. Obviamente, os custos continuaram altos. Surgiu então, na década de 70, outra filosofia, a “*CBM - Condition Based Maintenance*” que visa reduzir o custo de manutenção ao mesmo tempo em que supervisiona a condição de funcionamento dos equipamentos da planta de produção, prevenindo falhas e introduzindo o conceito de manutenção programada. O monitoramento das máquinas tem permitido detectar as falhas nos seus estágios iniciais, permitindo que a manutenção seja feita somente quando necessário, com planejamento, de forma a não interromper o processo além do estritamente necessário (Tavner & Penman, 1987). Um sistema de manutenção preditiva eficiente com monitoramento de faltas incipientes, também reduz o custo de manutenção evitando desmontagens para inspecionar peças e a substituição de peças pelo critério de tempo médio de uso.

Monitoramento da condição e proteção de sistemas têm funções relacionadas, no entanto, o primeiro é utilizado para prevenir falhas ou identificá-las nos estágios iniciais, enquanto a segunda tem ação posterior à ocorrência das faltas, sendo ainda dependente da intensidade com que ocorram.

Esta dissertação destina-se ao desenvolvimento de uma ferramenta que, posteriormente, poderá ser utilizada para diagnóstico de faltas on-line em motores de indução. O trabalho é parte integrante de um projeto de pesquisa mais amplo sobre Sistemas para Diagnóstico de Falhas em Máquinas Elétricas, financiado pela FAPEMIG. Consiste em obter

dados a partir de um simulador físico de faltas elétricas bem como identificar redes neurais aplicáveis ao problema de estudo de faltas em motores de indução, que, depois de adequadamente treinadas, serão capazes de avaliar os dados *off-line* recebidos e a partir daí, sinalizar a respeito da existência ou não de faltas, bem como indicar a fase onde estejam ocorrendo. Estas características situam o projeto como de monitoramento da condição.

1.2 - Formulação da Situação-Problema

Para obter-se o máximo benefício do capital investido nas plantas produtivas é necessário manter alta disponibilidade, o que só é possível com a melhoria dos projetos e do padrão de manutenção. A utilização de relés de proteção reduz a extensão dos danos quando ocorrem faltas, no entanto, apesar de serem efetivos e confiáveis, requerem um tempo, durante o qual poderá haver uma grave deterioração no isolamento da máquina. Como sempre é mais barato realizar pequenos reparos, a um custo mínimo, existe um incentivo econômico para se detectar defeitos em plantas antes que evoluam para falhas e permitam que a planta continue funcionando até um momento ideal para se fazer os reparos. Assim, as políticas de manutenção vêm se tornando cada vez mais importantes e passaram a ser vistas como fundamentais para que se possa garantir alta confiabilidade, segurança e baixos custos de produção. Isto só é possível se o gerenciamento dos complexos sistemas de produção forem automatizados, de forma a permitir a rápida detecção da situação de falta e também da mais provável causa do defeito. Pelo monitoramento das máquinas, é possível fazer desligamentos

programados, detectar faltas que, interrompidas em seus estágios iniciais, evitam perdas de tempo e dinheiro, com desmontagens, substituição de peças, perdas de produção, etc.

Atualmente, estima-se que motores de indução são responsáveis por mais de 80% do processo de conversão de energia nos setores industriais. Assim sendo, o monitoramento da condição de tais máquinas elétricas tem se tornado cada vez mais importante e necessário para melhorar sua confiabilidade naquelas funções em que são essenciais aos processos produtivos. Instrumentação e sistemas de diagnóstico têm sido desenvolvidos para examinar, regularmente, as grandezas elétricas e mecânicas associadas à máquina, de forma que as faltas possam ser corrigidas nos estágios iniciais, seja pela adoção de estratégias de operação ou manutenção programada com redução de custos de reparos.

1.3 - Objetivo do Estudo

O objetivo é desenvolver *um sistema para auxiliar na identificação da condição de funcionamento da máquina de indução trifásica sujeita à ocorrência de faltas no estator. Pretende-se que uma rede neural artificial efetue o diagnóstico de faltas elétricas em tais máquinas elétricas, a partir de dados monitorados e convertidos para valores por unidade (pu)*. O estudo estará restrito aos defeitos nos enrolamentos do estator de motores de indução trifásicos.

Os principais resultados esperados são: (a) estudar diferentes tipos e topologias de redes neurais, visando aplicá-las em processos de identificação das condições normais ou anormais de funcionamento de um motor de indução, com possibilidade de, no futuro, utilizar-se tal ferramenta na identificação de defeitos em rotores; (b) adquirir conhecimento e

domínio sobre modernas ferramentas tecnológicas como redes neurais; (c) tornar disponível, como resultado da pesquisa, uma ferramenta que possa, após as devidas adaptações, ser acoplada a uma placa de aquisição de dados para ser usada na detecção on-line de faltas em motores de indução; (d) conhecer o desempenho de diferentes redes neurais frente ao treinamento e identificação de defeitos a partir da mesma massa de dados.

1.4 - Justificativas

Nos dias atuais, as indústrias buscam maior produtividade, menor custo e maior confiabilidade no funcionamento de seus parques industriais. Este trabalho pode ser entendido como uma contribuição na área da manutenção preditiva, já que a detecção de faltas elétricas incipientes em motores de indução propicia economia de tempo, redução no custo de produção e no tempo de paralisação dos sistemas, colaborando para que seja estabelecida uma política eficiente de manutenção.

1.5 - Organização do texto

No capítulo 2 é apresentada a revisão bibliográfica, onde se procura levantar as várias metodologias, de diferentes tendências, estudadas por grupos científicos distintos, para realizar a detecção e o diagnóstico de faltas em sistemas físicos, plantas industriais ou equipamentos de forma geral; o estado-da-arte e as tendências dos vários centros de pesquisa,

além da identificação dos defeitos estatóricos que podem interferir no funcionamento de um motor de indução trifásico, de rotor em gaiola ou bobinado e sínteses dos artigos e estudos realizados nesta área.

No capítulo 3 são apresentados os procedimentos metodológicos aplicados à presente investigação, referentes à coleta e tratamento de dados, pressupostos e limitações gerais a serem considerados.

No capítulo 4 são apresentadas: as condições gerais, os pré-processamentos aos quais a base de dados obtida foi submetida; a metodologia de simulação e de avaliação de resultados, bem como a análise dos resultados obtidos nas diversas simulações feitas, buscando o tipo e configuração de rede que melhor desempenho apresentasse na identificação das condições de funcionamento da máquina de indução tomada como protótipo e ensaiada conforme apresentado no capítulo 3.

No capítulo 5 são apresentadas as conclusões gerais e as propostas de continuidade do estudo.

CAPÍTULO 2

REVISÃO DA LITERATURA

2.1 - Introdução

O objetivo deste capítulo é apresentar a revisão bibliográfica feita durante o preparo desta dissertação, buscando com isto, o estado-da-arte na atividade de monitorar máquinas elétricas de indução inseridas nos processos industriais.

A primeira parte deste capítulo apresenta a importância da detecção das faltas no contexto da manutenção. A segunda informa sobre as metodologias, de diferentes tendências, estudadas por grupos científicos distintos, para realizar a detecção e o diagnóstico de faltas em sistemas físicos, plantas industriais ou equipamentos de forma geral. A terceira, focando máquinas elétricas, identifica os defeitos que podem interferir no funcionamento da máquina de indução trifásica, de rotor em gaiola de esquilo ou bobinado, com especial atenção aos efeitos que afetam os enrolamentos estatóricos. Serão apresentadas, na quarta parte, sínteses dos artigos que apresentam alguma contribuição em termos de técnicas para detecção de faltas. Na quinta e última parte é apresentada a revisão da literatura referente à utilização de técnicas de inteligência computacional na detecção de faltas.

2.2 - A detecção das faltas no contexto da manutenção

Começemos por estabelecer, para este trabalho, o significado de dois termos amplamente utilizados na manutenção: falta e falha.

Para Isermann (1984), uma falta é “um desvio não permitido de uma propriedade característica a qual conduz à inabilidade de cumprir o propósito pretendido”. Faltas podem além disto, já estarem existindo ou aparecer em qualquer tempo; podem aparecer abruptamente ou muito lentamente. Sohlberg (1998) atribui às faltas lentas, origem em desgastes e àquelas de mudanças abruptas, origem em erros dentro do sistema de controle do processo.

Conforme You *et al* (1995) uma falta significa uma mudança no valor de um elemento em relação a seu valor nominal, podendo causar a falha de todo o sistema. Também define a falta *HARD*, como a falta catastrófica, produzida por um elemento que conduz a um curto-circuito ou a uma abertura do circuito. E define a falta *SOFT*, como a falta paramétrica que ocorre quando um elemento em falta desvia-se de seu valor nominal, mas permanece dentro de uma faixa de tolerância. Pode surgir devido a problemas de construção, idade ou efeitos parasitas.

Assim sendo, estaremos considerando, neste trabalho, que a falta em um equipamento conduz ao mau funcionamento, levando-o a outro estado diferente do estado de normalidade, que ocasiona uma falha do sistema no qual tal equipamento está inserido, trazendo prejuízos diversos, seja por perda de produção, custo de peças e mão-de-obra ou perda do próprio equipamento.

Quaisquer problemas mecânicos ou elétricos da máquina, bem como alimentação elétrica anormal, causam anomalias nas variáveis, permitindo que os estados anormais sejam percebidos, seja pela atuação da proteção seja pelas conseqüências desses estados anormais.

A proteção tem como objetivo interceptar as faltas durante os primeiros ciclos de ocorrência das mesmas e iniciar ações que interceptem os danos que possam estar sendo causados na máquina. De uma forma simples: um sinal é recebido por um transdutor que o converte em sinal elétrico que é traduzido pelo relé o qual, conforme sua programação de valores aceitáveis ou não, sinaliza a ocorrência do defeito e inicia uma ação, que, normalmente, consiste em retirar a máquina elétrica do circuito. Hoje, os relés estão muito sofisticados, absorvendo variadas funções programáveis que permitem cumprir suas funções com sucesso. Porém, para que a proteção sinalize a ocorrência de uma falta, a corrente de falta deve circular, trazendo danos como fundição de condutores e queima de isolamento. Além disto, a atuação da proteção de retirar a máquina do sistema de alimentação pode ter conseqüências sobre o sistema produtivo, no qual está inserido. Se a máquina está inserida em um processo vital para a produção, então haverá uma manutenção não programada do processo e, com certeza, o custo de paralisação da produção pode ser muito alto. Outro aspecto é que é preferível realizar pequenos reparos que podem ser feitos a um custo mínimo, do que ter de gastar grandes somas, imediatamente. Este é o grande incentivo econômico para se desejar detectar os defeitos da planta antes que se tornem falhas irreversíveis e continuar com o funcionamento até que seja encontrada uma boa oportunidade para realizar os reparos. A manutenção é, portanto, mais efetiva quando é planejada.

A possibilidade de se realizar manutenções programadas introduz a idéia e o interesse pelo monitoramento da condição. Embora proteção e monitoramento da condição tenham funções parecidas, são, em verdade, diferentes, tanto nos mecanismos e nas vantagens, quanto na forma de implementar. Enquanto o monitoramento tem como objetivo detectar a falta em

seus estágios iniciais, a proteção é essencialmente, retroativa, embora o monitoramento da condição possa ser estendido para prover proteção primária.

Com o constante avanço da tecnologia, os processos técnicos vêm crescendo em complexidade e exigindo cada vez mais confiabilidade e segurança. Além disto, a manutenção dos equipamentos elétricos é essencial para a operação segura e eficiente de qualquer indústria.

As estratégias de manutenção são variadas, conforme Pinto e Xavier (2003):

- a) Manutenção corretiva não planejada: correção de falha, após a ocorrência da mesma, de forma aleatória, sem planejamento, implicando em altos custos;
 - b) Manutenção corretiva planejada: manutenção feita por decisão gerencial, baseado em diagnóstico preditivo ou decisão de esperar pela quebra. Tem as vantagens de: ter um custo menor e permitir compatibilizar a intervenção com os interesses da produção, porém sua qualidade depende da qualidade da informação recebida de quem acompanha o estado de funcionamento do equipamento;
 - c) Manutenção preventiva: objetiva reduzir ou evitar as faltas, atuando no equipamento em intervalos de tempo pré-programados. O principal ponto negativo é a possibilidade de serem inseridos defeitos durante a manutenção, por imperícia humana, falta de peças adequadas para substituição das danificadas e contaminações de lubrificantes. Deve ser adotada caso a preditiva não esteja disponível;
 - d) Manutenção preditiva: consiste no acompanhamento sistemático das condições da máquina, promovendo recomendações para manutenção corretiva planejada dos equipamentos monitorados, apenas quando são atingidos os chamados “níveis de alarme”. Todas as obtenções de informações são feitas com o equipamento em pleno funcionamento, desta forma, é a manutenção que menos interfere com a produção. Reduz
-

as falhas classificadas como catastróficas e evita intervenções desnecessárias, permitindo a programação da manutenção, o que reduz os custos.

- e) Manutenção detectiva: surgiu na década de 90, objetivando descobrir as chamadas “falhas ocultas”, que, invisíveis às equipes de produção e manutenção, podem comprometer severamente o sistema produtivo. Tal detecção e retorno à condição de funcionamento também devem ser feitos, preferencialmente, mantendo o sistema em produção. Como exemplo, podemos citar o monitoramento dos sistemas de proteção e desligamento automático, quanto às possibilidades de não-atuação ou atuação indevida.
- f) Engenharia de manutenção: significa uma mudança cultural. Caracteriza-se por: procurar as causas reais da ocorrência das faltas; minimizar a ocorrência de problemas que se repetem com frequência; corrigir sistemáticas ou padrões inadequados; interferir nas compras vinculadas à manutenção e, muitas vezes, recomendar e orientar correções de projeto no processo produtivo. Contempla um novo paradigma de melhorar sempre, buscando os níveis de manutenção dos países do Primeiro Mundo.

A adoção de quaisquer das técnicas deve ser condicionada à importância que o equipamento tem para a planta na qual está inserido e a uma análise de custo x benefício.

O monitoramento da condição pode ser usado, tanto como ferramenta de diagnóstico quanto como técnica de detecção de faltas incipientes. Pode ser usado para simplesmente dar alarme como para mostrar as tendências e as leituras de desvios dos sensores, como também pode ser utilizado em um sistema de diagnóstico que, além das funções do sistema de monitoramento, utiliza regras básicas de inteligência computacional para relacionar as leituras dos sensores ao gerador de condições com um fator de confiança calculado. Este sistema gera mensagens para o operador da planta quanto à condição, um aconselhamento do que deve ser feito e as consequências caso não sejam tomadas as medidas indicadas. Basicamente, as diferenças entre um sistema de monitoramento e um sistema de diagnóstico incluem maior

rapidez e maior precisão na identificação de problemas, bem como a existência da regra básica de diagnóstico, que relaciona as leituras dos sensores e a condição de diagnóstico vinculada.

Um sistema de diagnóstico de faltas pode prevenir desligamentos desnecessários, reduzir custos de manutenção, evitando desmontagens para inspeção.

Tavner, Gaydon e Ward (1986) avaliam que o monitoramento da condição pode ser interessante ou não para uma máquina em particular, dependendo: do preço da máquina; do preço envolvido no conserto da mesma, caso ela sofra uma falta; da antecipação com que os avisos de uma falta iminente sejam dados; do custo e da confiabilidade do próprio esquema de monitoramento. O preço do conserto da máquina precisa ser considerado tendo em conta: o custo direto, o tempo e os custos de interrupção da produção envolvidos no reparo ou substituição, bem como a disponibilidade para substituição da máquina avariada. Os autores dão um exemplo muito interessante, considerando estes aspectos: se um monitoramento confiável, de alto valor, feito em uma máquina estratégica, pode detectar uma falta somente um segundo antes que a proteção possa operar, o monitoramento pode ser considerado como difícil de trazer qualquer benefício. Por outro lado, se o monitoramento puder, seguramente, dar uma hora ou mais de advertências sobre uma falta em tal máquina, este será de considerável valor e, se, dias ou meses de advertências puderem ser dados também sobre máquinas de menor valor, este poderá ser benéfico no planejamento da manutenção. Concluem que o monitoramento da condição de uma máquina elétrica produzirá benefícios se um sistema de monitoramento existente puder prover quantidade suficiente de mensagens de atenção de uma falta em andamento, tal que a máquina possa ser retirada de serviço e reparada a um custo que é substancialmente menor do que se as correntes de curto fluírem. O benefício líquido será positivo se os custos do sistema de monitoramento forem menores do que o economizado no reparo que se segue à ocorrência de faltas. Um monitoramento de

condição efetivo deve oferecer: medida de parâmetros selecionados para dar clara indicação de danos; processamento e avaliação do significado dessas medidas; ação a ser tomada em função dessas medidas.

Tavner e Penman (1987) fornecem um amplo panorama do monitoramento das máquinas elétricas, apresentando aspectos dos mecanismos das faltas, dos elementos de monitoramento, das diversas técnicas disponíveis para o monitoramento da condição das máquinas elétricas. Lembram também que muitas máquinas são incorretamente especificadas para a aplicação a que se destinam e é uma perda de tempo aplicar sofisticadas técnicas de monitoramento para uma máquina que esteja inadequada para seus propósitos, como é o caso de máquinas sub-dimensionadas ou com invólucro inadequado. A especificação deve levar em conta as condições mecânicas, elétricas e ambientais, sendo que também no caso de faltas, tais fatores influenciarão diretamente os mecanismos de faltas. Fazem algumas considerações sobre o monitoramento contínuo da condição de itens críticos da planta que podem conduzir a benefícios. Um destes seria o fato de que o operador da planta poderá receber, continuamente, informações atualizadas sobre o desempenho da máquina, o que poderá ajudá-lo a melhorar a disponibilidade operacional e a eficiência da planta. O monitoramento da condição deverá dar informação relevante para atendimento às funções de manutenção e operacional.

Atribuem à função monitoramento as seguintes tarefas: (a) transdução do sinal primário, (b) aquisição, (c) processamento de dados e (d) diagnóstico.

Em Natarajan (1989), as limitações das políticas tradicionais de manutenção são citadas e o monitoramento das faltas incipientes é apontado como forma de evitar custos e problemas associados com faltas inesperadas.

No artigo de Siyambalapitiya (1990), encontra-se uma análise econômica e financeira da viabilidade dos sistemas de monitoramento de máquinas elétricas de indução. Em dois estudos de caso, os custos do projeto, construção, montagem e operação dos sistemas de

monitoramento de condições são discutidos e estimativas reais são feitas em custos da época. Estes custos são comparados aos custos de falhas inesperadas, produção perdida, reparos e perdas de produção. São utilizadas funções de probabilidade para a ocorrência de faltas e uma função para a medida da capacidade do sistema de monitoramento para detectar faltas. São apresentados dois estudos de caso.

Cash, Habetler e Kliman (1998) levantam aspectos muito interessantes: (a) frente às novas e promissoras tecnologias, os benefícios potenciais devem ser considerados, tendo em conta os custos de desenvolvimento e implantação. O investimento deve ser economicamente prudente; (b) máquinas pequenas podem ou não ser críticas para os processos principais e ainda assim merecerem a política de substituição no caso de falhas; (c) grandes máquinas com danos elétricos internos podem evoluir rapidamente e terem pequena correlação com a perda de produção; (d) os modernos métodos de monitoramento da condição das máquinas elétricas devem evitar o uso de sensores adicionais como medidores de temperatura e vibração, por exemplo, adotando a extração das informações a partir de tensões e correntes terminais.

Sohlberg (1998) divide os procedimentos para tratar faltas abruptas em três diferentes fases: (a) detecção: consiste em identificar a ocorrência de uma falta real, frente a distúrbios; (b) diagnóstico: consiste em localizar e determinar o tipo de falta, identificando precisamente, se possível, quais componentes do processo foram afetados; (c) impacto da ocorrência da falta dentro do sistema em que está inserido: uma decisão é tomada quanto à continuidade de funcionamento ou parada para manutenção. Pode-se utilizar a estatística para avaliação das variáveis quanto aos seus valores médios, magnitudes e variâncias admissíveis. A este módulo pode ser incorporado um módulo de análise de risco.

Isermann (1984) descreve as etapas do processo de supervisão da mesma forma que Sohlberg (1998), porém após o diagnóstico da falta, considera como passo seguinte, a avaliação da falta, determinando como ela afetará o processo. Uma vez conhecidos os efeitos

da falta, a decisão sobre a ação a ser tomada pode ser feita: no caso de falta tolerável, a operação pode continuar ou pode ser definida uma mudança de operação. Se a falta for intolerável, a operação deve parar e a falta deve ser eliminada imediatamente.

2.3 - Métodos para detecção e diagnóstico de faltas em processos produtivos

Em termos genéricos, os métodos de detecção de faltas podem ser baseados em: a) modelos de processos; b) métodos de estimação de estado ou parâmetros; c) métodos de decisão estatística; d) métodos de reconhecimento de padrões; e) métodos que utilizam sistemas de diagnóstico analógico e f) sistemas especialistas. Também como característica, tais métodos devem ser sensíveis ao surgimento das faltas, mas insensíveis a outras mudanças como ruídos, variações de sinais, erros de modelagem, etc.

2.3.1 - O uso de modelos

Como ponto de partida, há que se considerar que a modelagem de sistemas tem se desenvolvido paralelamente ao conhecimento humano. O propósito dos modelos é permitir a realização de estudos sobre os sistemas, objetos ou fenômenos que apresentem dificuldades na análise natural, dificuldades de acesso, medição ou mesmo altos custos ou riscos envolvidos. Construído um modelo, a simulação de tal modelo torna-se ferramenta para o estudo do comportamento do sistema.

Kourosch e Chin (1991) classificam os modelos para diagnósticos em: mecanicistas (baseados em modelos) ou empíricos (baseados em conhecimento). Como mecanicistas, definem aqueles modelos que são derivados da física do processo. Têm fundamentação analítica e utilizam técnicas de identificação paramétricas para diagnóstico. São modelos de estado dinâmico que assumem conhecimento completo das características dos ruídos das medições e representam os processos de faltas por suas variáveis de estado (ver Isermann, 1984). Também afirmam que a maioria dos sistemas mecânicos apresenta alta complexidade para uma perfeita modelagem mecanicista.

Os modelos empíricos, por outro lado, são desenvolvidos sobre heurísticas estabelecidas experimentalmente ou definidas em relação a indícios de faltas. Como estes modelos não necessitam de compreensão detalhada do processo, são muito utilizados no monitoramento de sistemas complexos. Podem ser citados, como pertencentes a este grupo, os modelos que se utilizam de inteligência computacional e que são a base para sistemas especialistas de diagnóstico.

O processo de detecção antecipado de faltas é um pré-requisito essencial para o desenvolvimento de um sistema de supervisão automático, o que pode ser feito pela verificação de variáveis em relação aos valores considerados como normais ou toleráveis.

Isermann (1984) apresenta um modelo matemático para estudo de um processo generalizado e seus sinais:

$$Y = F \{U, N, \theta, x\} \quad (1)$$

Onde:

N = sinais de distúrbios não mensuráveis;

θ = parâmetros não mensuráveis do processo;

x = variáveis internas de estado parcialmente mensuráveis e parcialmente não mensuráveis;

Y = sinais de saída;

U = sinais de entrada.

Para o autor, as seguintes grandezas são básicas para caracterização dos métodos de detecção das faltas:

- a) sinais mensuráveis (U, Y);
- b) variáveis de estado não mensuráveis;
- c) parâmetros não mensuráveis do processo;
- d) grandezas características não mensuráveis $\eta = g(U, Y, \theta)$.

Os métodos baseados em sinais mensuráveis utilizam os sinais de entrada $U(t)$ e de saída $Y(t)$ para monitorar mudanças no processo. Uma das técnicas é monitorar o sinal de saída $Y(t)$, tal que, caso sejam medidos valores acima do previsto ou abaixo do definido como mínimo, um sinal é emitido. Análises do sinal de saída também podem dar informações sobre o estado interno do processo.

Os métodos de estimação requerem modelos matemáticos do processo. Sohlberg (1998) lembra que as vantagens e desvantagens de cada método devem ser consideradas em cada situação. Um exemplo é a estimação on-line de parâmetros realizada durante a ocorrência de faltas, que pode estimar valores errados e conduzir a decisões errôneas na manutenção.

A detecção das faltas baseado na estimação das variáveis de estado é feita na reconstrução ou estimação dessas variáveis a partir de sinais mensuráveis pelo uso de um conhecido modelo de processo.

Para estimação, um processo dinâmico, em que não se utilizam observadores não lineares:

$$X(t) = f \{U, Y, t\} \quad (2)$$

deve ser linearizado em relação a um ponto de operação, desta forma o processo deve operar próximo ao ponto onde a linearização foi feita, porque o modelo só será válido nas redondezas desse ponto de operação.

Fazendo a mudança de representação de estado e considerando que as variáveis de estado de interesse $x_i(t)$ são elementos do vetor de estado $x(t)$:

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (3)$$

$$y(t) = Cx(t) \quad (4)$$

Considere-se que: $x(t)$ é o vetor de estado;

A é a matriz de transição de estado do sistema;

B é a matriz de entrada;

y é o vetor de saída;

C é a matriz de transformação;

$y = \Delta Y$ – mudanças nas variáveis de saída Y;

$u = \Delta U$ – mudanças nas variáveis de entrada U;

$x = \Delta X$ – mudanças nas variáveis de estado X.

Para reconstruir estes estados a partir de sinais de entrada e saída mensuráveis, deve-se selecionar uma matriz H de realimentação:

$$\hat{x}(t) = A\hat{x}(t) + Bu(t) + H[y(t) - C\hat{x}(t)] \quad (5)$$

Mudanças nos processos, causadas por ruídos ou ação de atuadores podem ser modelados por $v(t)$:

$$\dot{x}(t) = Ax(t) + Bu(t) + Fv(t) + v(t) \quad (6)$$

Ruídos nas medidas por $n(t)$ e mudanças nos sensores por $\mu(t)$ em:

$$y(t) = Cx(t) + n(t) + \mu(t) \quad (7)$$

Ainda, conforme Isermann (1984), mudanças abruptas de estados e das saídas podem ser modeladas se $v(t)$ e $n(t)$ são considerados como impulsos delta ou funções quadradas. Essas mudanças podem ser detectadas através da geração de resíduos, pelo uso de filtros de Kalman-Bucy:

$$\tilde{y}(t) = y(t) - C\hat{x}(t) \quad (8)$$

Isermann (1984) ainda cita alguns métodos especiais de teste para embasamento das decisões de faltas: (a) filtros sensíveis à faltas: abordagem determinística onde a matriz de realimentação H é escolhida de modo que os modos de falta se manifestem como resíduos em uma direção ou plano fixo; (b) o teste de resíduos do filtro normal de Kalman; (c) um banco finito de filtros de Kalman – consiste em avaliar a resposta do sistema a múltiplos testes de modelos, com hipóteses de faltas incluídas; (d) teste da relação de semelhanças generalizadas – consiste em estabelecer correlações entre os resíduos observados e as respostas de filtros

devidos a certas faltas. Todos estes métodos requerem conhecimento exato dos parâmetros do processo e dos sinais que o influenciam.

Sohlberg (1998) faz referência ao método estatístico da “Soma de Resíduos Quadrados Ponderados” (WSSR), que complementado pelo teste da relação de semelhanças generalizadas (GLR) se aplica à detecção de faltas. Outro método citado é o do banco de filtros de Kalman, porém, a fraqueza deste método é basear-se no conhecimento preciso dos tipos de faltas que podem afetar o sistema, uma vez que é muito difícil conhecer todos os tipos de faltas que podem ocorrer.

Os métodos de detecção de faltas baseados na estimação de parâmetros de processos enfrentam, muitas vezes, dificuldades adicionais pelo fato de que os parâmetros, constantes ou coeficientes dependentes do tempo, utilizados na modelagem do processo, raramente se relacionam com facilidade aos coeficientes do processo físico. Desta forma, também deverá ser considerada a forma como os defeitos afetam os coeficientes físicos os quais, por sua vez, afetam os parâmetros do processo. Como coeficientes do processo físico, podemos citar: comprimento, massa, velocidade, viscosidade, resistências, capacitâncias, etc.

Para os casos em que os coeficientes do processo físico que indicam a ocorrência das faltas, não sejam medidos diretamente, Isermann (1984) apresenta um procedimento para detectar as faltas a partir de mudanças nos parâmetros do modelo do processo: a) estabelecer as equações do processo em função das variáveis de entrada e saída; b) determinar o relacionamento entre os parâmetros do modelo e os coeficientes do processo físico; c) fazer a estimação dos parâmetros do processo como resultado dos sinais de entrada e saída; d) calcular os coeficientes do processo e determinar suas alterações; e) fazer um catálogo das faltas nas quais fica estabelecido um relacionamento entre o processo da falta e as mudanças nos coeficientes do processo físico. Assim é possível identificar as faltas, inclusive por reconhecimento de padrões.

Os métodos de detecção baseados em grandezas mensuráveis utilizam-se das variáveis mensuráveis de entrada (“U”) e de saída (“Y”) para obter informações a respeito do estado interno de grandes plantas. São exemplos de grandezas mensuráveis: rendimento da máquina; consumo de combustível por produção unitária ou tempo; consumo de óleo em motores a combustão, compressores, etc., por unidade de produção ou tempo; desgaste de ferramentas, e de componentes como rolamentos, por unidade de produção ou tempo. Essas grandezas características podem apontar problemas como contaminação, desgastes excessivos, vazamentos, etc.

Basicamente, pode-se observar que os métodos de detecção de faltas por estimação de parâmetros ou de estados monitoram as variáveis de saída do sistema em relação aos valores previstos. Quaisquer deficiências no sistema geram desvios em relação aos valores previstos, gerando resíduos, sinais de erro ou mudanças nos parâmetros. As decisões são então tomadas em função do diagnóstico feito sobre esses desvios, utilizando técnicas no domínio do tempo ou da frequência.

Na construção de modelos são utilizadas duas abordagens: a) simulação de circuito equivalente, para construir modelos não paramétricos; b) identificação de sistemas para estimar modelos paramétricos. A primeira opção permite construir modelos de faltas sem a necessidade de provocar faltas físicas, mas requer grande esforço computacional para gerar os modelos e efetuar as comparações. A segunda pode oferecer um grande conjunto de soluções, limitadas em relação ao sistema em análise e às suas modalidades de operação.

Ostovic (1988) apresenta em seu artigo, uma abordagem simplificada para a modelagem dos circuitos equivalentes magnéticos das máquinas de indução, buscando melhoria na velocidade e precisão dos cálculos, pelo fato de eliminar as iterações na solução do sistema de equações algébricas. Assegura que o método é, em essência, um tipo de análise

por elementos finitos e que pode ser tão preciso quanto necessário. No entanto, uma redução de precisão significa economia de tempo de processamento e memória. Os resultados obtidos são representados em sistema coordenado espaço-tempo de todas as grandezas eletromagnéticas. As técnicas de modelagem utilizadas consideram todas as particularidades da máquina, como tipo e conexão dos enrolamentos, tipo de ranhuras e saturação do núcleo de ferro, tanto em regime quanto em transitórios. As matrizes são amplamente utilizadas na teoria de circuito equivalente magnético, por conectarem as grandezas do circuito magnético, os fluxos nos ramos e as forças magnetomotrizes, com as correntes e fluxos por fase da máquina. Utiliza o método de potencial nodal. No capítulo 5 de seu livro, Ostovic aplica sua metodologia no exame de falhas em máquinas polifásicas, com os enrolamentos conectados de diferentes formas e alimentadas tanto em corrente como em tensão.

Borgard, Olsson e Lorenz (1995) apresentam um modelo de 5^a. ordem da máquina de indução, contendo a dinâmica de 3 subsistemas: elétrico, magnético e mecânico. Através da análise destes três subsistemas e da seleção de taxas de amostragens apropriadas para cada subsistema, os parâmetros dinâmicos da máquina podem ser identificados com maior precisão, evitando os problemas relacionados à resolução finita de conversores analógico-digitais com a taxa de amostragem. Foi encontrado um modelo o qual, estando em um ponto de operação correto permite que seus parâmetros físicos variem com o espaço de solução.

Sullivan e Sanders (1995) desenvolveram um novo modelo em π para a máquina de indução, baseado em um modelo magnético de um par de dentes rotor-estator, com dois elementos com relutância não linear representando a saturação no rotor e no estator. Tal modelo é compatível com a transformação de Blondel-Park, guarda maior correspondência com o modelo físico e é mais fácil utilizá-lo na simulação e controle.

A modelagem também pode ser utilizada na identificação das faltas, incorporando o tipo e magnitude da falta ao modelo (Sohlberg, 1998). Desta forma, o comportamento do

processo pode ser analisado por simulação e fornecer informações ao operador sobre o estado do processo, o que favorecerá a tomada de decisão. Outra forma é utilizar as propriedades estatísticas do processo, como magnitude, valores médios e variância das variáveis mais importantes e verificar tais valores. Aconselha-se incluir dentre os procedimentos de identificação das faltas, um módulo de análise de risco para avaliar o risco de continuidade da operação, ou das atitudes tomadas.

Os chamados modelos de caixa branca, cinzenta e preta sintetizam princípios de modelagem.

Os modelos de caixa branca são construídos a partir do conhecimento completo do processo, aplicando-se também como modelo de simulação. Tais modelos são ideais para aqueles processos descritos de forma completa e precisa por equações matemáticas, como equações diferenciais, equações algébricas, relacionamentos lógicos e tipos similares de equações. Um exemplo é a utilização das leis de Kirchoff para modelar circuitos elétricos.

Nos modelos de caixa preta, o conhecimento é obtido a partir de dados experimentais, já que se aplicam a processos considerados como completamente desconhecidos e em relação aos quais não se necessita de qualquer modelo que reflita a estrutura física do processo. O processo é descrito usando grupos de modelos padrões e estimação de parâmetros desconhecidos, a partir de experimentos. Exemplos de aplicação deste tipo de modelagem são os sistemas complicados e incertos, como os econômicos, sociais, ecológicos e processos industriais que tenham o conhecimento incompleto sobre sua estrutura.

Os modelos de caixa cinzenta são utilizados naqueles processos, sobre os quais há apenas conhecimento incompleto em relação à sua estrutura. O conhecimento físico disponível é incompleto e incerto, influenciado por distúrbios. Estas características situam os modelos de caixa cinzenta em posição intermediária em relação aos demais modelos: enquanto os de caixa preta necessitam somente identificação e os de caixa branca apenas

modelagem, os de caixa cinzenta requerem tanto modelagem quanto identificação. Pode-se dizer que requerem parte de conhecimento do modelo e parte de dados experimentais. Além disto, é possível ver o processo como um conjunto de submodelos ou modelos internos, o que permite otimizar o desempenho do sistema. Na construção de um modelo cinzento é necessário ter em mente as várias atividades independentes, porém, altamente vinculadas necessárias: modelagem básica, expansão da modelagem, experimentação, identificação, análise e avaliação do modelo.

Quanto à flexibilidade, os métodos baseados em modelos são pouco flexíveis, na medida em que mudanças posteriores são mais difíceis de serem incorporadas.

Isermann (1984) apresenta uma estrutura generalizada dos métodos de detecção de faltas, reproduzida na FIG. 1, que concentra, objetivamente, as similaridades entre os vários modelos.

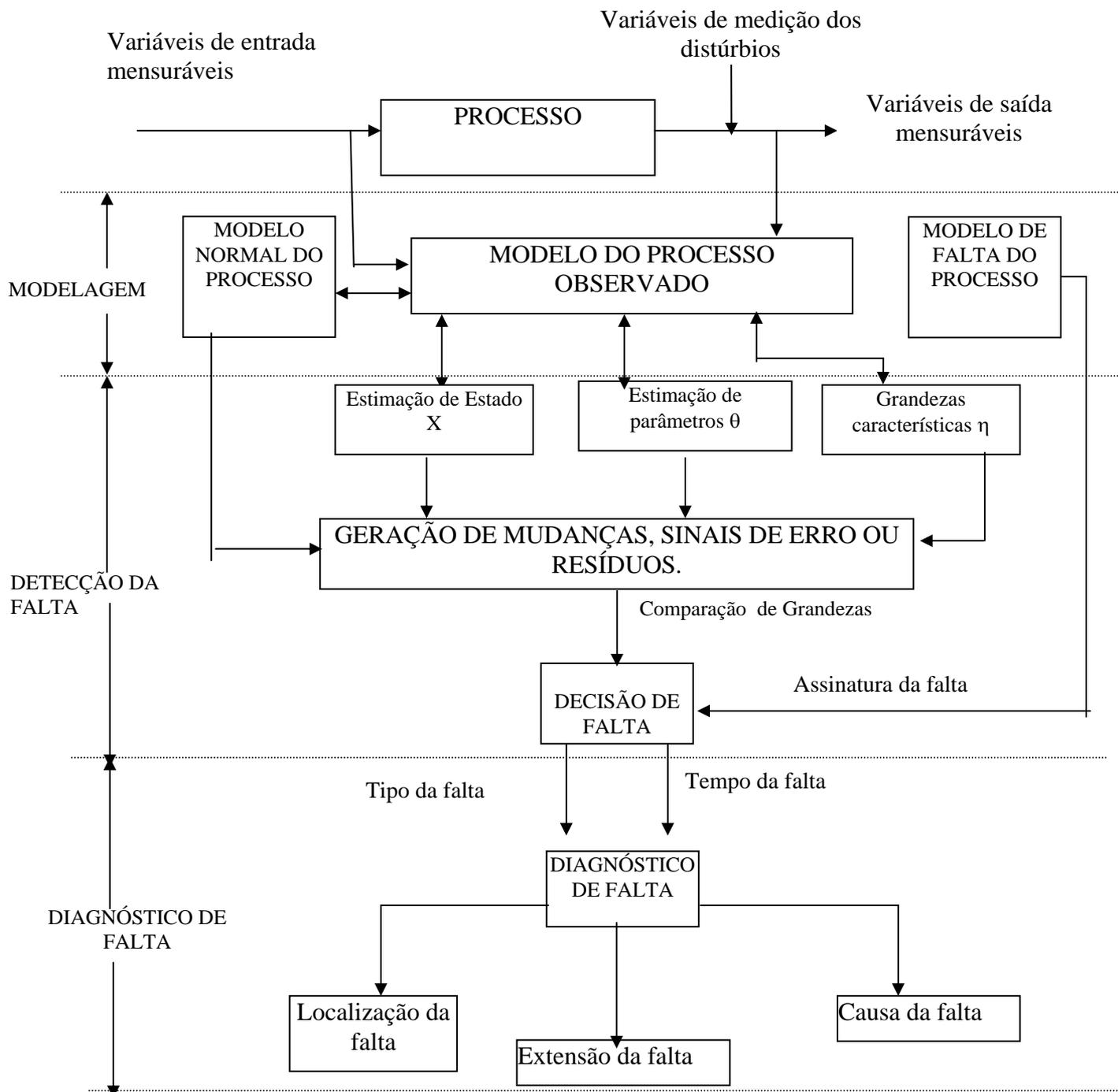


FIGURA 1 – Estrutura generalizada dos métodos de detecção de falhas baseados em modelos de processos, grandezas mensuráveis e não mensuráveis.

FONTE: Isermann, 1984, p. 391.

2.3.2 - Métodos que utilizam reconhecimento de padrões

Enquanto a detecção e o diagnóstico de faltas baseadas no monitoramento de parâmetros e variáveis requer modelos teóricos precisos e métodos de estimação eficientes, os métodos de reconhecimento de padrões não requerem necessariamente qualquer modelo matemático, se estiverem baseados em dados obtidos experimentalmente, para efeito de simulações e análises.

Tais métodos podem ser encarados como um mapeamento do espaço de medição no espaço de decisão, ocorrendo uma classificação de acordo com os dados medidos.

Segundo Kourosh e Chin (1991) as técnicas de classificação estatística de padrões foram propostas para diagnóstico de faltas como forma de enfrentar a variabilidade dos sinais de faltas, através da criação de “regiões de decisão” associadas com as faltas individuais. Um fator a ser considerado é o alto custo de se fazer a aquisição experimental de sinais de faltas necessários à formação de regiões de decisão confiáveis.

Sorsa, Koivo e Koivisto (1991) dividem o reconhecimento tradicional de padrões e a classificação em 3 estágios: medição, extração de características e classificação. À etapa de medição sucede-se a extração de informação redundante do vetor de características. O vetor de características é classificado em uma ou mais classes, sendo este o principal objetivo do reconhecimento de padrões. Se a detecção de faltas está associada ao diagnóstico, algumas classes podem ser referentes à condição normal de funcionamento e outras às condições de defeito.

Os métodos e reconhecimento de padrões, conforme Sood et al. (1985) podem ser classificados em uma de duas categorias: (a) estatístico ou de decisão teórica e (b) estrutural ou sintático.

O desenvolvimento de regras de decisão para a classificação de faltas poderá se tornar algo complexo ou simples em função do acerto no cálculo das características.

Basicamente, os métodos de reconhecimento de padrões vêm suprir a deficiência dos métodos baseados em modelos, no tocante à necessidade de incluir, a priori, todas as situações em estudo e de estarem relacionados a um ponto de operação, em especial.

Os métodos de reconhecimento requerem, no entanto, uma amostragem representativa de dados, os quais têm importância fundamental nesses métodos.

O esforço computacional, segundo os autores, é surpreendentemente menor do que naqueles métodos baseados em modelos, especialmente naqueles que requerem modelos não-lineares.

Os métodos de reconhecimento de padrões podem ser estatísticos ou baseados em inteligência computacional. Podemos citar, como exemplo de utilização de padrões estatísticos, o trabalho de Sood, Friedlander e Fahs (1985), que desenvolveram e aplicaram, com sucesso, três metodologias para localização de faltas em cilindros de motores multicilindrados: (a) correlação estatística; (b) medida de similaridades e (c) classificação de Bayes. Concluíram que o método de similaridades é mais simples de se implementar computacionalmente em relação ao método de reconhecimento de padrões, no entanto, a precisão do segundo é superior ao do primeiro.

Sood, Fahs e Henein (1985), em outro artigo, utilizam a abordagem de estimação de parâmetros e comparam esta técnica com as de similaridades e de reconhecimento de padrões. Concluem que no caso da estimação de parâmetros, quatro equações diferenciais devem ser resolvidas a cada iteração, sendo que são necessárias cerca de 15 iterações para encontrar uma solução satisfatória, onerando computacionalmente, o método. Apresentam como vantagens, o fato do método não requerer dados de treinamento e possuir uma baixa taxa de erro com valor intermediário entre a medida da similaridade e o método baseado em Bayes.

Dentro ainda de técnicas de estimação de parâmetros, cabe ressaltar o uso das redes neurais como ferramenta de inteligência computacional.

Visando imitar a facilidade que os seres humanos têm de reconhecer padrões e efetuar classificações, surgiram as redes neurais, capazes de fazer um mapeamento transparente entre o espaço de medição e o espaço de decisão e aprender através de exemplos de treinamento.

As redes neurais artificiais oferecem simplicidade e alta velocidade operacional na solução de complexos problemas de diagnóstico, mas, para que sejam adequadas ao propósito de diagnóstico, deve-se escolher adequadamente e normalizar as entradas.

(Filippetti, Franceschini, & Tassoni, 1993) apresentam argumentos que reforçam as vantagens de se utilizar redes neurais artificiais: (a) os modelos desenvolvidos para a análise de faltas rotóricas, no caso de estudos de máquinas elétricas de indução, utilizando outras abordagens, necessitam de várias hipóteses que podem conduzir a resultados inaceitáveis; (b) quando tais modelos necessitam de dados de projeto, os mesmos podem não estar disponíveis, como no caso de máquinas em operação há muito tempo; (c) uma vez que as entradas da rede estejam escolhidas corretamente e o treinamento tenha sido feito baseado em dados com e sem defeitos, as interconexões conterão as características do sistema físico, evitando-se os problemas relativos à modelagem matemática.

Kovács (1996) define dentro do universo de sistemas de controle de processos, as funções essenciais voltadas ao processo como temporização, sequenciamento, sincronização e o intertravamento de eventos e as funções essenciais voltadas ao operador: monitoração, supervisão, detecção de falhas e o processamento de alarmes. Acrescenta que nas funções voltadas ao processo, a aplicação de redes neurais tem pequeno ganho se comparadas às técnicas convencionais, no entanto, no caso de funções voltadas ao operador, a utilização dessa técnica tem sido bem explorada pelos pesquisadores, já que as redes neurais são classificadas como excelentes detectores de padrões. Em especial, neste tipo de aplicação, a

rede neural deve ser treinada a partir de um conjunto de dados gerados através de um modelo de simulação de faltas do sistema. Frequentemente, a produção desse modelo pode ser mais complexa que o treinamento da rede neural. Uma dificuldade adicional é relacionar com exatidão, os parâmetros do modelo e os coeficientes físicos como: comprimento, massa, velocidade, viscosidade, resistências, capacitâncias, dentre outros.

2.3.3 - Sistemas Especialistas

Durante a segunda metade da década de 1960, surgiram os trabalhos sobre os sistemas que hoje conhecemos como especialistas, por resolverem problemas que normalmente são solucionados por “especialistas” humanos. Há grandes dificuldades para se extrair informações de especialistas humanos, seja pela pouca clareza, seja pela falta de detalhes de mais baixo nível, ou de informações estatísticas. O conhecimento do especialista na aplicação deve ser traduzido de forma substancial em uma base, criada da forma mais eficiente possível. Os sistemas especialistas devem induzir automaticamente suas próprias regras através de exemplos de problemas e soluções, por isto, são programas complexos, baseados em regras, que utilizam a inteligência computacional e precisam explorar um ou mais mecanismos de raciocínio, para aplicar o conhecimento na solução dos problemas. Também é necessário que as pessoas interajam com o sistema especialista de forma fácil e eficaz, para isto ele deve estar capacitado para explicar seu raciocínio, adquirir novos conhecimentos e modificar os anteriores. O sistema especialista deve ser capaz de identificar pontos fracos na base de conhecimentos e solicitar complementação. Rich e Knight (1993) exemplificam alguns dos sistemas especialistas em utilização hoje: (a) XCON – usado para configurar os sistemas DEC

VAX; (b) o PROSPECTOR que aconselha sobre a exploração de minerais; (c) o DESIGN ADVISOR que avalia projetos de chips; (d) o MYCIN que recomenda terapias para pacientes com infecções bacteriológicas; (e) o MOLE usado para diagnóstico de problemas de classificação heurística, como é o caso do diagnóstico de doenças, problemas com motores de carros ou em siderúrgicas e ineficiências em usinas termelétricas.

Gonzalez *et al.* (1986) utilizaram o programa chamado PDS – Sistema de Diagnóstico de Processos – para desenvolver um sistema especialista para realizar diagnósticos *on-line* de turbo-geradores a vapor. Tal ferramenta permite que profissionais outros que os programadores criem o sistema através de menus direcionados iterativamente. O processo de ensinar ao programa é conhecido como “engenharia do conhecimento”, através de “regras básicas” que contém o conhecimento do diagnóstico. Tais regras têm a forma IF...THEN. A regra básica é utilizada por uma “máquina de inferência” que simula o processo do pensamento humano, para realizar o diagnóstico, apresentando conclusões com um fator de confiança associado. Isto é, reconhece quando uma regra, em particular, pode ser aplicada. Tão logo a máquina de inferência identifica a situação que se apresenta como a parte da regra correspondente ao IF, a ação vinculada ao THEN é tomada. O PDS é a ferramenta de inteligência computacional que define um conjunto genérico de conceitos tais como sensores, regras e hipóteses para representar o conhecimento especialista, a ser utilizado em diagnósticos *on-line*. As regras básicas são construídas a partir desses conceitos do PDS e são compostas por nós, regras e funções: (a) o nó de mais baixo nível é um sensor ou nó de entrada; (b) o próximo nível de nós é a hipótese, à qual se vincula um nível de confiança; (c) a regra conecta o nó sensor ao nó hipótese e mapeia a leitura do sensor a um nível de confiança que se refere à hipótese – verdadeira ou falsa. Isto é feito através de uma função que assinala o fator de confiança ao correspondente sensor de leitura; (d) o nó de mais alto nível é o nó da condição, que representa a condição monitorada do equipamento a ser comunicada ao

operador. Tal nó é mantido por uma ou mais hipóteses através de regras contendo funções ponderadas.

Milne (1987) afirma que o ramo de diagnóstico é um dos maiores domínios de aplicação de sistemas especialistas e que o uso de regras para fazer o mapeamento entre as possíveis causas, as entradas de um sistema e as possíveis faltas têm sido substituídas por novas técnicas que permitem construir regras a partir de informação probabilística, aprendizado por experiência ou exemplos. Descreve a máquina de inferência como um mecanismo simples em que cada uma de um conjunto de regras é experimentada por vez. Apresenta os diferentes níveis de representação do conhecimento e afirma que diferentes sistemas de diagnóstico podem ser desenvolvidos a partir desses níveis de conhecimento. Sendo o conhecimento básico necessário para diagnóstico, o conjunto de defeitos e relações entre observações e defeitos, o sistema pode obtê-lo diretamente ou a partir de modelos em profundidade ou conhecimento causal. A detecção de uma situação anormal começa com um comportamento que apresenta um desvio em relação ao esperado ou desejado. O sistema deverá gerar hipóteses acerca da causa do defeito, hipóteses essas que podem ser obtidas a partir da observação de defeitos específicos já indubitavelmente vinculados às hipóteses. Neste caso, fazem uso do que se chama de pedaços de conhecimento pré-compilados, os quais relacionam as observações comportamentais (valores testados, sinais, sintomas, etc.) a um certo número de hipóteses, que podem ser testadas como forma de se confirmar ou rejeitar as hipóteses. Estes testes também serão gerados pelo sistema especialista. Se este conhecimento não estiver disponível para o sistema, é comum a utilização de modelos de faltas, cujos parâmetros e características gerais são utilizados na identificação de eventuais defeitos não mapeados pelo sistema especialista.

Milne também apresenta um diagrama dos quatro níveis de representação do conhecimento. São eles: o estrutural ou de conectividade, o comportamental, o funcional e a

comparação de padrões. O sistema diagnóstico pode utilizar quaisquer desses níveis como nível de entrada (conhecimento estrutural primário, comportamental, funcional ou padrões compilados), permitindo derivar informação para qualquer outro nível superior de representação do modelo, saindo com a solução do problema. Alternativa é entrar no sistema especialista, no nível de comparação de padrões compilados e sair com um sistema baseado em regras. Os sistemas de conhecimento compilados não utilizam nenhum modelo estrutural em profundidade são, por isto, freqüentemente, referidos como sistemas de conhecimento de superfície ou rasos. Os sistemas de diagnóstico são, em sua maioria, construídos usando a abordagem baseada em regras, na qual as regras existem para concluir sobre possíveis faltas. As regras são organizadas tal que uma ou mais hipóteses intermediárias são calculadas e depois, combinadas para a análise final.

O tipo mais simples de conhecimento é aquele utilizado em sistemas diagnósticos estruturais, por isto pode gerar sistemas rápidos, utilizados para isolar regiões faltosas diretamente, a partir da descrição do sistema. Os conhecimentos sobre o comportamento e função podem ser usados para, então identificar o problema exato, eliminando possíveis hipóteses candidatas.

Naqueles sistemas em que a inferência é feita sobre dados reais trabalhados vindos do ambiente ou a partir de dados medidos diretamente, é necessário um primeiro passo, durante o qual os dados numéricos são traduzidos em valores qualitativos –fase da redução- durante a qual são utilizados critérios específicos. Tais informações qualitativas darão a direção de que questões são fundamentais.

A utilização de quaisquer dos modelos - comportamental, funcional ou causal - permite a representação dos mecanismos de representação do conhecimento, para uma grande quantidade de informação, permitindo construir sistemas de diagnóstico potentes e complexos. Eles são utilizados, primariamente como mecanismos de simulação. A introdução

de modelos de faltas torna-se crucial para confirmar hipóteses de faltas. Pode-se acoplar aos modelos em profundidade, conhecimento experimental, o que pode ser visto como primeiro passo na direção do aprendizado ou auto-aperfeiçoamento do sistema. O aprendizado tem sido focado sobre refinamento de regras ou sobre a incorporação de conhecimento experimental.

Para Protopapas et al (1990) a única diferença entre os vários sistemas especialistas é o método de raciocínio e representação simbólica do conhecimento na base de conhecimento. Um software eficiente dá rapidez e simplicidade aos métodos de representação do conhecimento. Além disto, a facilidade de incorporar conhecimento na forma de linguagem natural, permite que um usuário, possa utilizar o sistema e alimentá-lo, mesmo se não possuir maiores conhecimentos de informática. Afirma que a confiabilidade de uma planta industrial depende do programa de manutenção das máquinas. Sistemas especialistas com muito conhecimento no reparo e manutenção de máquinas elétricas, atuam com eficiência, segurança e rapidez, no apoio aos engenheiros de manutenção.

Kourosch e Chin (1991) apresentam um novo método para diagnóstico de faltas para lidar com a variabilidade dos sinais de faltas em sistemas complexos. Trata-se de um método não parametrizado de classificação de padrões para diagnóstico de faltas, que não requer conhecimento sobre a estrutura estatística do sistema. Utiliza uma função para discriminação linear dos padrões de classificação com o uso de realimentação para estimação ou atualização do modelo de diagnóstico. Essa função de discriminação permite conhecer a capacidade de diagnosticar condições de funcionamento do sistema, detectando redundâncias entre medidas, a qual pode ser usada para seleção e otimização de sensores. Assim, um sistema é diagnosticável se cada falta é identificada com exclusividade.

Sharaf, Clark e Wright (1992) apresentam um programa de diagnóstico on-line baseado em regras de reconhecimento de padrões. O sistema de diagnóstico é baseado na comparação de padrões no domínio da frequência, estando composto pelos módulos: (a) de

aquisição de grandezas como tensões, correntes e velocidade obtidas utilizando análises no domínio da frequência; (b) de tratamento estatístico; (c) de reconhecimento de padrões e identificação que emprega a Transformada Rápida de Fourier (FFT). Este módulo utiliza sinais adquiridos e filtrados para reconhecer padrões de falta para desenvolvimento de regras usando gráficos no domínio da frequência.

2.4 - Os defeitos em máquinas elétricas e suas causas

A minimização dos problemas em máquinas elétricas requer compreensão dos padrões de projeto, construção e instalação das máquinas. O funcionamento depende da correta instalação, comissionamento e da implementação das rotinas operacionais apropriadas e procedimentos de manutenção. Assim, Clark, McShane & Wakeley (1987) apresentam um artigo sobre os fatores construtivos e de projeto que podem contribuir para a ocorrência de faltas.

Falhas no isolamento podem ocorrer devido à vibração mecânica, variações de tensão, calor, idade, erosão, defeitos durante a instalação (Chow & Yee, 1991a) ou construtivos, tais como vácuos e corpos estranhos embebidos na parede do isolamento, penetração de óleo ou metal no isolamento. Também podem ocorrer por manutenção constante de temperaturas acima daquela ditada pela classe de isolamento da máquina.

Bonnett e Soukup (1992) admitem que a maioria das faltas é causada pela combinação de vários estresses que atuam sobre os principais componentes da máquina elétrica. Sobre o estator, identificam os grupos de *estresse* caracterizados pela origem de suas causas: térmica, mecânica, ambiental e elétrica. Para o rotor, individualmente, podem ser apontadas ainda as

causas de origem eletromagnética ou dinâmica. Além desses aspectos, outros como o tipo de máquina, o ambiente onde elas trabalham e o suprimento de energia, podem definir os mecanismos pelos quais as máquinas elétricas falham. A identificação dos primeiros sinais dessas faltas permite que o monitoramento torne-se possível e benéfico.

Os defeitos rotóricos podem ser o resultado das seguintes forças e condições: força dinâmica desbalanceada; vibração torcional e conjugados transitórios; problemas de fundição, soldagem, usinagem e encaixes radiais, axiais e outros; forças magnéticas causadas por fluxos de dispersão vibrando a duas vezes a frequência da corrente rotórica; força magnética causada por excentricidade do entreferro; forças centrífugas; *estresse* térmico causado por aquecimento dos anéis terminais, por temperatura diferencial em barras durante a partida (efeito pelicular) ou por crescimento axial das barras; forças axiais causadas pela inclinação das barras rotóricas.

Em relação aos defeitos rotóricos, foram identificadas várias bibliografias, dentre as quais citamos: Williamson et al. (1982), os quais afirmam que se a velocidade do motor flutua por causa da mudança de carga ou como parte do ciclo de carga normal, pode ocorrer falta por fadiga nas juntas entre barras e anel nos rotores em gaiola de esquilo; Kerszenbaum e Landy (1984) mostram a importância de considerar as correntes laminares interbarras em grandes motores de indução de rotor em gaiola com barras ou anéis quebrados. Mostra os efeitos dessas correntes tangenciais sobre a produção de vibrações axiais, devido à presença de fluxo axial, com possível queima de laminações na região próxima à falta. Desenvolveram uma expressão que descreve a distribuição das correntes dentro de um rotor em gaiola com uma barra quebrada, as quais variam com o escorregamento.

Bonnett e Soukup (1992) apresentam análises de faltas nos enrolamentos, identificando os vários tipos de modos e padrões de faltas e relacionam-nos às prováveis causas específicas de tais faltas. Citam cinco áreas chaves para diagnosticar a falta. São elas:

o modo, o padrão e a aparência da falta, características da aplicação em si e a história de manutenção.

Os modos de falta também podem ser agrupados nos 5 grupos: falta entre espiras, entre bobinas, circuito aberto, entre fases e fase-terra.

Afirmam que, na análise de faltas é difícil determinar que condições mencionadas anteriormente iniciaram o problema e qual foi o resultado provocado. Citam um motor enrolado e colocado em funcionamento com um curto parcial entre espiras, dentro de uma bobina. Como esta condição progride, gera-se um aquecimento excessivo dentro da bobina em curto, resultando na deterioração do isolamento e eventualmente em um aterramento parcial da bobina na ranhura. Dependendo do tipo de proteção do motor, ele pode continuar a rodar; o aumento do calor gerado na área provoca destruição do isolamento. Neste ponto, uma falta fase-fase direta ou para a terra ocorre e o motor é rapidamente retirado de funcionamento. Neste caso, como em outros, a inspeção pode revelar todos os 5 modos de falta, mas a condição de falta entre espiras, foi o problema inicial e os outros resultaram dele. Uma falta entre espiras é usualmente muito difícil de reconhecer, devido à natureza destrutiva das condições finais de falta.

Os padrões de falta são estritamente relacionados aos modos de falta, mas podem ser classificados em: simétrico, fase simples, assimétrico com aterramento e assimétrico sem aterramento.

Para se realizar um diagnóstico preciso é necessário relacionar modos e padrões de faltas, além disto, a aparência geral do motor dá indícios das possíveis causas das faltas. Deve-se observar a situação de limpeza dos enrolamentos, presença de materiais estranhos, umidade, sinais de sobreaquecimento, sinais de travamento do rotor, partes mecânicas perdidas, localização da falta, inclusive em relação às fases. Também há que se considerar as condições gerais de operação no momento da falta. Alguns aspectos a serem observados

quanto às condições de operação são: carga característica do acionamento, pulsações ou ciclicidade das cargas, aceleração da carga, resistência do enrolamento, tensões desbalanceadas, transitórios de tensão, outros motores que falharam na mesma aplicação, tipo de aceleração e de proteção, condições ambientais. A história de manutenção da máquina também dá indicações da causa do problema.

Nas seções 2.4.1 a 2.4.4 serão apresentados alguns aspectos de como estes grupos de estresses se relacionam, como afetam a vida da máquina e contribuem para a ocorrência de faltas prematuras.

2.4.1 ESTRESSE DE ORIGEM TÉRMICA

1. Envelhecimento: o efeito causado pelo envelhecimento térmico do enrolamento é o de conferir vulnerabilidade ao sistema de isolamento quanto a outros fatores que produzem faltas, como o *estresse* mecânico e ambiental. A perda de integridade física, caracterizada por ressecamento e aspecto quebradiço, pode levar a faltas tão só pelo estado em que se encontram. Deve-se observar que altas temperaturas de operação podem ser geradas pela circulação de corrente elétrica acima dos valores nominais previstos para a máquina, por longos períodos. A vida térmica pode ser aumentada, reduzindo-se a temperatura de operação ou aumentando-se a classe de isolamento.
 2. Sobrecarga térmica por:
 - Variações de tensão: o desempenho dos motores e a temperatura do enrolamento ficam afetados por variações fora da faixa de $\pm 10\%$ da tensão nominal, prevista para o equipamento.
 - Tensões de fase desbalanceadas: um pequeno desbalanceamento causa um incremento grande de temperatura no enrolamento: estima-se que a cada 3,5% de desbalanceamento, corresponderá um aumento de temperatura de 25% na fase com a mais alta corrente.
 - Sobrecarga: dentro de certos limites, pode-se estimar que a temperatura do enrolamento aumentará com o quadrado da carga. O recomendado é que o motor funcione abaixo dos limites normais de temperatura previstos para sua classe de isolamento ou que a máquina seja superdimensionada em termos de classe de isolamento.
 - Partidas repetitivas: dependendo do tipo do motor, a corrente na partida pode variar de 5 a 8 vezes a corrente que circula em condições de carga nominal. Se um motor está
-

sujeito a repetidas partidas em um pequeno período de tempo, a temperatura aumentará rapidamente.

- Obstrução da ventilação: o calor gerado no rotor e estator é dissipado por condução, convecção e radiação. Qualquer coisa que obstrua o fluxo de ar ou impeça a radiação do calor das partes do motor, causará um incremento na temperatura dos enrolamentos. Também é importante assegurar que o motor seja mantido limpo por fora e por dentro, para que o fluxo de ar não seja reduzido.
- Temperatura ambiente: os motores padrões são projetados para operar em uma temperatura ambiente máxima de 40° C. Operações acima desse limite podem reduzir as horas de vida do isolamento.

2.4.2 ESTRESSE DE ORIGEM ELÉTRICA

- Dielétrico: os materiais isolantes suportam até um certo nível de tensão. Esta tolerância é ditada pelos materiais que entram na composição dos mesmos. A especificação do tipo de material isolante deve ser feita de acordo com a aplicação a que se destina o motor, para assegurar tempo de vida adequado. As tensões elétricas que caracterizam este tipo de estresse podem ser classificadas em: entre fases, entre espiras e entre espira e massa.
 - Tracking: em motores operando acima de 600 V, o mecanismo desta falta inicia-se através de uma combinação entre umidade e materiais estranhos que constrói uma ponte de alta resistência entre o fio de cobre e a terra. Devido à diferença de potencial, uma pequena corrente circula para a terra, causando crescentes queimaduras no sistema de
-

isolamento. Com a evolução desta condição, o isolamento para a terra deteriora-se ocasionando a falta.

- Corona: este fenômeno caracteriza-se por descargas localizadas, resultantes da ionização gasosa transitória em um sistema de isolamento operando acima de 5 kV, onde a tensão aplicada exceda um certo valor considerado crítico. São tipos básicos de descargas: as descargas internas que ocorrem nas cavidades do dielétrico; as descargas que podem ocorrer na superfície das bobinas; os pontos de descarga que ocorrem em um forte campo elétrico ao redor de uma ponta (poder das pontas). Os fatores que afetam a descarga por corona incluem: frequência, espessura do dielétrico, material, vácuos, nível de tensão, geometria, umidade, estresse mecânico e temperatura. O mecanismo de falta é o aquecimento, desgaste ou reação química, resultando em uma deterioração do isolamento do enrolamento.

- Condições transitórias de tensão: a exposição dos motores a condições transitórias de tensão causa falhas no isolamento dos enrolamentos, porque o *estresse* elétrico não é uniformemente distribuído através dos enrolamentos.

Essas tensões transitórias podem ser causadas por: curtos-circuitos; altas oscilações de tensão causada por repetitivos religamentos; oscilações de tensão ou ressonâncias causadas por interrupção de corrente em circuitos indutivos, quando a energia do campo magnético não é zero; rápida transferência de barramentos, gerando um vetor diferença entre a tensão do barramento reserva e a tensão em decaimento; abertura e fechamento de disjuntores, gerando ondas de impulso que dão origem a tensões irregulares; chaveamento de capacitores, gerando surtos em função das capacitâncias utilizadas; perfurações no isolamento em outros pontos do sistema de potência que podem criar, em sistemas de alta tensão, surtos de tensão que excedem o triplo das tensões fase-terra normais em um sistema não solidamente aterrado; surtos de tensão gerados por raios através de descarga

direta ou indução; acionamentos de frequência variável cujos efeitos surgem nos enrolamentos como faltas espira-a-espira ou espira-terra; rápidas variações de tensão originadas no chaveamento de inversores.

2.4.3 ESTRESSE DE ORIGEM MECÂNICA

- Partidas repetitivas: além dos efeitos mencionados anteriormente, devido às elevadas correntes na partida, surgem forças de origem eletromagnética que podem causar deformação nas cabeças das bobinas do estator e do rotor, provocando a expansão ou contração do sistema de isolamento, cujos materiais, em longo prazo, podem se deteriorar. Outro efeito é o de vibração, que causa afrouxamento dos enrolamentos, degradação das escovas, do isolamento e dos mancais de rolamentos.

O primeiro objetivo deve ser o de proteger a cabeça das bobinas contra grandes forças durante transitórios de carga e o segundo, amortecer as barras condutoras contra forças menores durante o funcionamento contínuo, em regime. Faltas podem ocorrer sobre as cabeças dos enrolamentos quando a braçadeira da estrutura afrouxa como um resultado de uma sucessão de sobrecargas pouco usuais ou por causa de um extenso período de funcionamento contínuo. Em alguns casos, o isolamento das cabeças das bobinas torna-se gasto. Em máquinas maiores, falhas nos enrolamentos por fadiga podem ocorrer quando o enrolamento fica frouxo o suficiente para permitir um movimento significativo do condutor durante operação normal, na partida ou re-chaveamento. Indicações do problema são: incremento na vibração do enrolamento e a possibilidade de descargas para a terra.

A vibração excessiva ou a ocorrência de defeitos construtivos podem ser os responsáveis pela ocorrência de curtos entre as laminações no núcleo do estator, fazendo com que as correntes circulantes no núcleo assumam valores elevados, chegando a derreter chapas.

- Movimento de bobinas: este efeito ocorre com maior intensidade em grandes máquinas de alta velocidade, do que em pequenas máquinas de baixa velocidade. Caracteriza-se pelo fato de que, durante a partida, a força que surge nas bobinas, produzida pela corrente no enrolamento estático, causa vibração nas bobinas com uma frequência de duas vezes a frequência de linha, criando movimentos nas direções radial e tangencial. Este movimento pode causar severos danos ao isolamento das bobinas e aos condutores de cobre.
 - Golpes do rotor: existem casos em que o rotor pode golpear o estator, como por exemplo em situações de falhas nos rolamentos, deflexão do eixo e desalinhamento entre estator e rotor. Se o contato entre estator e rotor ocorrer na partida, a força do rotor pode fazer com que as laminações do estator perfurem o isolamento da bobina, aterrando-a. Se o contato ocorre quando o motor está rodando na velocidade nominal, o resultado é um prematuro aterramento das bobinas na ranhura do estator, causado pelo excessivo calor gerado neste ponto de contato.
 - Outras causas: pesos que fazem o balanceamento do rotor soltarem-se, laminações quebradas por fadiga, lâminas do ventilador soltas, porcas e parafusos perdidos; penetração de partículas estranhas através do sistema de ventilação; defeitos no rotor.
-

2.4.4 ESTRESSE DE ORIGEM AMBIENTAL (contaminação)

Uma das coisas mais importantes é manter o motor limpo e seco, interna e externamente. Como isto não é fácil, o ideal é selecionar um invólucro e/ou sistema de isolamento que permita uma maior proteção contra contaminação. A presença de material estranho pode causar: redução na dissipação de calor, causando aumento da temperatura de operação; conduzindo a falta prematura nos enrolamentos por estresse localizado e também por curtos-circuitos.

Os principais efeitos sobre os mecanismos de falta derivam da temperatura, umidade e presença de poeira. Altas temperaturas podem causar deterioração do isolamento; baixas temperaturas causam congelamento; alta umidade pode causar condensação e a baixa umidade, ressecamento dos solventes usados no isolamento. A presença de poeira pode contaminar o isolamento ou os componentes mecânicos. Os efeitos térmicos e de contaminação podem ser conjugados: a máquina pode estar funcionando fora dos limites especificados em sua classe de isolamento, seja por problemas de refrigeração, condições ambientais ou operação fora dos limites nominais. A máquina pode estar operando em um ambiente empoeirado, seja pelo processo industrial onde a máquina está inserida, seja devido à poeira de carbono produzida pelo desgaste das escovas. Se a poeira entrar no circuito de refrigeração principal, pode contaminar enrolamentos, escovas e conexões elétricas, causando deterioração da integridade do isolamento. Também pode haver sobreaquecimento e danos mecânicos. Qualquer desses pode conduzir à condensação da umidade sobre o isolamento elétrico e conexões, baixando a resistência de isolamento.

Na TAB. 1 é apresentada uma síntese proposta por Bonnett e Soukup (1992), relacionando as causas de faltas aos padrões de queima dos enrolamentos.

TABELA 1

Síntese de metodologia relacionando as várias causas de faltas aos padrões de queima dos enrolamentos.

PADRÃO	POSSÍVEIS CAUSAS	APARÊNCIA DO ENROLAMENTO	MODO DE FALTA
Padrão de queima simétrica com todas as fases sobreaquecidas	<ul style="list-style-type: none"> Alta ou baixa tensão; Carga excessiva; Grande número de partidas; Necessidade de ventilação própria; Condição ambiental elevada; Defeitos no rotor; Falta completa dos enrolamentos; 	As bobinas com sinais de sobreaquecimento evidenciam descoloração e falta no isolamento, dependendo da quantidade de calor.	As faltas reais ocorrem usualmente devido a uma combinação de curtos e aberturas, podendo o enrolamento também ser aterrado devido a extremo aquecimento na ranhura do estator ou condutores.
Padrão de queima monofásico	<ul style="list-style-type: none"> Controle ou suprimento de potência monofásico; Abertura dos fios ou condutores do enrolamento; Conexão inadequada; Fonte de tensão desbalanceada; 	Dependendo do tipo de conexão em estrela ou triângulo, uma ou duas fases podem apresentar pontos com temperatura elevada que podem conduzir a faltas devido a um curto entre espiras.	Se a causa é interna ao enrolamento, a detecção de fase ou fases não aquecidas pode indicar circuito aberto.
Padrão de queima assimétrico onde o enrolamento é aterrado. Obs.: dependendo do tipo de proteção, a falta para a terra pode ser o tipo mais destrutivo de falta. Pode haver destruição do enrolamento, danos na laminação, elevados riscos de choque e perigo para o pessoal da operação.	<ul style="list-style-type: none"> Atrito entre o rotor e as laminações do estator durante a partida ou funcionamento; Danos no isolamento, ranhuras, anéis terminais ou condutores; Transitórios de tensão, surtos por chaveamento ou raios; Contaminação por poeira, produtos químicos ou materiais estranhos; Caminho de baixa tensão ou deterioração 	A aparência da falta no enrolamento normalmente se limita a manchas específicas nas ranhuras do estator.	O modo de falta primário é bobina-terra. Contudo, podem existir sinais de curtos entre espiras e entre fases.

	<p>por corona no isolamento;</p> <ul style="list-style-type: none"> • Sobreaquecimento nas ranhuras do estator devido a corrente excessiva ou precária dissipação de calor; • Movimento de bobinas dentro das ranhuras; • Vibrações das cabeças das bobinas. 		
Padrões de queima assimétrico sem aterramento.	<p>Muitos dos itens já citados responsáveis por faltas envolvendo a terra podem também causar falta espira-a-espira. O fator determinante está vinculado à fraqueza ou resistência do sistema de isolamento: em um ambiente empoeirado, a falta será no ponto de maior fraqueza do isolamento. Se existe movimento anterior da bobina nas espiras terminais, resultantes de algum dano, o modo de falta poderá ser espira-a-espira. Se o isolamento da ranhura do estator foi enfraquecido pelo mesmo movimento da bobina, então o modo de falta pode ser fase-fase ou bobina-bobina. As causas são as mesmas dos padrões assimétricos com aterramento, exceto para atritos entre estator e rotor.</p>	<p>A aparência evidencia curtos entre espiras e aberturas, sem aparência geral de aquecimento. Contudo existirão sinais de grande aquecimento na área vizinha à de falta.</p>	<p>Freqüentemente, uma falta fase-fase ocorre e a proteção retira o motor de funcionamento.</p>

FONTE: Bonnett e Soukup (1992).

2.5 - Técnicas utilizadas na detecção de faltas em máquinas elétricas, baseadas em monitoramento de grandezas elétricas, magnéticas e mecânicas

A base de qualquer método de detecção confiável é a compreensão do comportamento elétrico e eletromagnético da máquina sujeita às condições de falta.

Os mais freqüentes tipos de problemas que podem ser esperados em um motor de indução são relacionados à alimentação do motor, estrutura mecânica, enrolamentos estatóricos e rotóricos. Todo problema modifica as variáveis elétricas, magnéticas e mecânicas do motor e o especialista humano, ou não, deve ser capaz de identificar o problema.

Como visto no item 2.2, os métodos baseados em sinais mensuráveis utilizam os sinais de entrada $U(t)$ ou de saída $Y(t)$ para obter informações a respeito do estado interno de grandes plantas e para monitorar mudanças nos processos.

A seguir serão apresentadas algumas sínteses de artigos que reproduzem este método.

(Williamson et al., 1982) afirmam que faltas em desenvolvimento durante o funcionamento são mais difíceis de serem identificadas e que havia, na literatura, métodos de detecção baseados nas flutuações da velocidade, ruídos acústicos e espalhamento do fluxo.

(Cameron et al., 1986) reportam um estudo para detectar excentricidades estáticas e dinâmicas no entreferro de grandes motores de indução trifásicos, utilizando método não invasivo: monitoramento da vibração estatórica e da corrente de linha. Conjugando os métodos, os autores asseguram identificação única do evento em andamento, além de requererem equipamentos simples como acelerômetros e transformadores de corrente, facilmente instalados no ambiente industrial.

(Filippetti et al.,1988) afirmam que um dos mais adequados parâmetros para identificar qualquer problema é, sem dúvida, o campo magnético, mas a instalação de bobinas de pesquisa ou outros transdutores magnéticos nem sempre é prático ou econômico, então diferentes possibilidades devem ser consideradas. Uma delas diz respeito ao uso das correntes de linha que são mais acessíveis pelas técnicas de medição não invasivas, porém contêm menor quantidade de informação do que a densidade de fluxo magnético. Além disto, o monitoramento por correntes estatóricas conduz a indicações úteis sobre o estado de saúde da máquina e sinalizam alterações nas condições de operação, permitindo ações que evitem maiores danos, no caso de faltas. Citam a detecção de algumas outras variáveis encontradas na literatura: campos magnéticos axiais e tangenciais, campo magnético no eixo, conjugado instantâneo, correntes estatóricas, vibrações, velocidade rotórica e posição. No entanto, segundo os autores, apesar do processamento dessas variáveis produzir diagnósticos não ambíguos dos eventos, estes estudos são orientados para a solução de problemas específicos.

Siyambalapatiya, (1990), cita as técnicas de: (a) análise química para detecção da degradação térmica da isolação; (b) monitoramento da corrente estatórica para detecção de faltas rotóricas; (c) técnicas de monitoramento on-line baseado na medição de fluxos axiais; (d) medição direta das variáveis rotóricas.

Embora o monitoramento térmico e de vibrações tenha sido utilizado por décadas, grande parte das pesquisas recentes foi direcionada para o monitoramento elétrico da máquina com ênfase na inspeção da corrente de fase da máquina (Schoen et al., 1995). Em especial, o espectro de corrente estatórica tem sido usado para perceber faltas rotóricas associadas com barras rotóricas quebradas e desbalanceamento mecânico. Todas as técnicas disponíveis requerem um certo grau de especialização para distinguir-se entre a condição normal de operação e uma falta potencial. Isto ocorre porque os componentes espectrais, tanto vibração quanto corrente, podem resultar de fontes diversas, incluindo-se aí aquelas relacionadas às

condições normais de operação. Outro aspecto a considerar é a grande quantidade de harmônicos existentes devido ao projeto e construção da máquina e às variações na carga acionada. Muitos dos harmônicos podem ser causados por um rotor que esteja ovalizado, com porosidades ou bolhas nas peças fundidas, projeto de ranhuras, etc. O problema pode ser avaliado considerando-se que variações cíclicas de carga e desbalanceamentos mecânicos podem ter exatamente o mesmo efeito sobre o espectro de corrente estatórica resultante de uma barra rotórica quebrada ou um eixo inclinado.

Cash, Habetler e Kliman (1998) referem-se a alguns dos métodos off-line para se detectar falhas incipientes no isolamento: (a) testes de resistência; (b) testes do fator de dissipação; (c) testes de surto; (d) testes da constante dielétrica; (e) testes de alta frequência; (f) testes de tensão suportável e (g) inspeção visual. E a alguns métodos on-line: vetores de Park, descargas parciais, componentes simétricas, potência na fase, comparações de modelos em tempo real.

Na bibliografia preliminar foram encontradas algumas técnicas de modelagem e/ou simulação de faltas em motores de indução como em:

Krause e Thomas (1965) utilizaram a simulação em computador analógico para estudar o desempenho das máquinas de indução simétricas, tanto bifásicas como trifásicas, durante operações balanceadas e desbalanceadas. A representação foi desenvolvida de forma a simular a máquina simétrica em qualquer sistema de referência. A variação das indutâncias mútuas com o ângulo de deslocamento do rotor θ_r , faz com que coeficientes variantes no tempo apareçam nas equações de tensão. Esta característica foi eliminada pela adequada mudança de variáveis, vinculada ao relacionamento trigonométrico entre os eixos do estator, do rotor e de um terceiro conjunto ortogonal (eixos “d-q”) definido como rotativo na velocidade angular elétrica arbitrária “ ω_r ”. Os modos de operação estudados foram: (a) condições balanceadas; (b) tensões estatóricas desbalanceadas; (c) resistores rotóricos

desiguais; (d) uma combinação de tensões estatóricas desequilibradas e resistores rotóricos desiguais; (e) abertura e fechamento das fases do estator.

Williamson et al. (1982) fazem uma análise baseada na teoria de campos rotativos, usando um modelo de malhas do rotor. A abordagem é geral e permite adaptação a qualquer tipo de falta rotórica, simétrica ou não. Os resultados obtidos para um grande motor industrial mostram que uma falta em barra única pode ser difícil de detectar em termos de componentes da corrente estatórica de frequência diferente daquela da alimentação ou do componente de dupla frequência de escorregamento do conjugado pulsativo, devido à pequena amplitude desses componentes comparados à frequência de alimentação e ao conjugado de regime.

Williamson e Mirzoian (1985) apresentam um método que permite prever o desempenho de um motor em gaiola no caso de faltas no enrolamento estatórico. O trabalho é baseado na versão generalizada da análise harmônica. As impedâncias de acoplamento são calculadas relacionando a tensão aplicada através de qualquer enrolamento à corrente que flui nos vários circuitos dentro da máquina. Um enrolamento é definido como um conjunto de bobinas conectadas em série. São apresentadas análises de faltas fase-neutro e fase-fase, sendo que a abordagem é a de análise por malhas. Segundo os autores, o método poderá ser de valia no ajuste de equipamentos de proteção que utilizam a componente de sequência negativa da corrente do estator para detectar faltas em andamento.

Williamson e Abdel-Magied (1987) apresentam um modelo matricial para estudo de assimetrias rotóricas em uma máquina de indução de dupla gaiola. Como também neste caso a interação entre estator e rotor com assimetrias é essencialmente a mesma que ocorre em uma máquina de indução com gaiola simples em falta, as correntes estatóricas também neste caso terão componentes trifásicas nas frequências “ f ” e “ $(1-2s)f$ ”, sendo “ s ” o escorregamento. Cada uma destas componentes criará campos em relação ao rotor. A equação de conjugado foi então desenvolvida tendo em vista estas duas componentes de força magnetomotriz do estator.

O método adotado para máquinas de gaiola simples, foi o de introduzir cada falta pela combinação e deleção de linhas e colunas da matriz de impedâncias que definem o sistema. Cada falta reduz a dimensão da matriz em um. No caso de máquinas com dupla gaiola, a técnica usada foi a de alterar as resistências representativas das barras fraturadas ou segmentos dos anéis terminais para um grande múltiplo de seu valor normal. Foram definidas as impedâncias de transferência de seqüência positiva e negativa em função de grandezas medidas nos terminais da máquina durante operação normal. A diferença entre as duas impedâncias permite discriminar a localização das faltas nas gaiolas externa e interna. Neste trabalho foi mostrado que as faltas na gaiola externa podem ser caracterizadas por sua resposta à excitação de seqüência negativa do estator, enquanto as faltas na gaiola interna podem ser identificadas pelos mesmos métodos usados para gaiola simples.

Maier (1992) apresenta um método para identificação de faltas para a terra, faltas na fase e curtos-circuitos de no mínimo 5% do enrolamento. O princípio básico do monitoramento do motor antes de partir é aplicar nos enrolamentos do motor, tensões de pulso e analisar a resposta. Quando o motor está em movimento o motor é monitorado através de um microprocessador. O método não é sensível a mudanças na carga nem a pequenas alterações no suprimento de tensão.

Apesar do objetivo ser a detecção de faltas no estator da máquina, o estudo da bibliografia mostrou que a corrente estatórica pode ser utilizada para detectar uma grande variedade de faltas. Em relação a estes sinais de corrente, as principais características de procedimento para um diagnóstico efetivo de faltas rotóricas são: processamento do espectro, identificação dos componentes da banda de corrente e correlação entre os componentes de corrente e a severidade das faltas.

Os artigos apresentados, a seguir, apresentam experiências, resultados e conclusões, considerados como de interesse, por fornecer reflexões e idéias sobre como realizar experimentos.

Albright (1971), descreve um efetivo método de detectar faltas entre espiras nos enrolamentos rotóricos de turbo-geradores a vapor. Consiste em medir a taxa de mudança da densidade de fluxo do entreferro através de uma pequena bobina de pesquisa, posicionada próximo à superfície rotórica. A saída em tensão é gravada por um filme Polaroid na tela de um osciloscópio e pode ser aí analisada. Os dados obtidos identificam as bobinas, as ranhuras, onde há espiras em curto e provê boa indicação do número de espiras em curto e a ocorrência de mais de um evento de faltas.

Em (Hargis, Gaydon, & Kamash, 1982) os vários defeitos rotóricos são estudados. Idealmente, um resultado instantâneo poderia estar disponível, sem interrupção do funcionamento da máquina e além disto, permitir a diferenciação entre defeitos provocados por barras rotóricas fraturadas ou com resistências anômalas, anisotropia magnética e flutuações de carga. O artigo apresenta várias técnicas práticas para determinação da condição do rotor de uma máquina de indução. As irregularidades eletromagnéticas podem ser quantificadas tanto por análise de corrente, quanto por velocidade. Ambas podem ser usadas para monitorar o progresso de um defeito, como forma de se reforçarem os diagnósticos, já que, individualmente, podem estar sujeitas às faltas. A análise de vibração não produz resultados quantitativos, mas é útil como parte de um programa de monitoramento. Os autores também asseguram que a análise da corrente é a técnica mais viável onde a máquina esteja inacessível ou nos casos em que não estejam disponíveis transdutores de vibrações ou geradores de pulso.

Reason (1987), em seu artigo, apresenta um programa conhecido como Motormonitor desenvolvido pela Entek Scientific Corp., de Cincinnati, Ohio. Tal software especialista

utiliza-se de alterações no espectro de frequências da corrente de operação para identificar problemas como: barras quebradas, anéis rachados, juntas de alta resistência, porosidades e bolhas no alumínio, conexões pobremente revestidas em cobre, irregularidades no entreferro, desbalanceamento magnético.

(Kliman et al., 1988) utilizam a análise do espectro de altas e baixas frequências, da corrente de linha para pesquisar harmônicos que caracterizem defeitos como: barras quebradas, desalinhamento da gaiola, variação na condutividade e problemas de orientação magnética das laminações. A corrente de linha e o fluxo axial externo são funções no tempo que serão transformadas por FFT em espectro. O algoritmo de decisão identifica a presença ou ausência de harmônicos que são típicos em descontinuidades nas gaiolas e geram mensagens.

(Filippetti et al., 1988) apresentam a representação do conhecimento do desempenho de uma máquina de indução com o propósito de identificar operações incorretas e faltas. Tal conhecimento é representado sob forma de árvore semântica, onde características diferenciadas conduzem à identificação do provável diagnóstico. As regras da árvore constituem-se em estratégias de inferência para um sistema especialista de diagnóstico da condição da máquina de indução. Utilizam como variáveis de entrada, correntes e tensões instantâneas de linha e a velocidade rotórica. Sugerem ainda a análise do conteúdo em termos de componentes de seqüência negativa para sinalizar desequilíbrio de tensões e, assim, a ocorrência de uma falta na máquina. Recomendam também um modelo para faltas estatóricas, no qual são feitas tantas equações estatóricas quantas sejam as frequências que surjam como resultado do defeito.

Natarajan (1989), apresenta uma técnica de monitoramento de falhas não aplicável a máquinas conectadas em triângulo ou em estrela não aterrada. O método baseia-se na medição das correntes estatóricas desequilibradas, usando para isto um TC de janela nas máquinas em

estrela aterrada. Ao invés de analisar as componentes espectrais do sinal, as magnitudes absolutas são usadas para identificar as faltas.

Quanto ao artigo de Trutt (1989), os procedimentos apresentados são baseados nos princípios de análise harmônica generalizada aplicada a motores de indução de rotores enrolados. A abordagem matricial permite o cálculo de parâmetros de operação com um número arbitrário, localização e tipo de simultâneas faltas estatóricas e/ou rotóricas. O desenvolvimento teórico utiliza a análise em séries de Fourier da força magnetomotriz retangular produzida por uma corrente cossenoidal nas bobinas do estator. A partir de tal análise, obtém-se as equações das impedâncias próprias e mútuas entre as bobinas do estator para as variações harmônicas e entre estas e as bobinas do rotor. As informações requeridas para cálculo das indutâncias devem ser obtidas junto aos fabricantes. A maior dificuldade na aplicação do método proposto é a precisão requerida na obtenção de dados de entrada confiáveis, particularmente as impedâncias de dispersão, próprias e mútuas entre bobinas e/ou enrolamentos.

Kliman e Stein (1990) revêm algumas das técnicas de detecção de faltas que utilizam a corrente estatórica para detectar problemas na carga, no eixo ou na alimentação, já que o monitoramento da corrente pode ser implementado sem maiores despesas. Examinam várias técnicas para detecção de faltas estatóricas e rotóricas em motores de indução, pela análise da corrente. Dentre estas, pode-se citar a detecção de corrente de falta para terra, a presença de correntes de seqüência negativa, a análise do espectro de freqüências da corrente, ajustes no circuito equivalente tradicional, etc. Conclui que várias tecnologias deverão ser integradas em um sistema de diagnóstico buscando aumento de sensibilidade e a resolução eficiente de ambigüidades.

Uma técnica diferente demonstrada por Vas (1993) consiste no monitoramento dos fluxos de dispersão axiais para identificar vários tipos de assimetrias. Baseia-se no fato de que

o conteúdo harmônico das correntes do estator e rotor reflete-se no conteúdo harmônico dos fluxos concatenados axiais do estator e rotor. Isto ocorre quando uma máquina não é simétrica, como no caso de faltas desequilibradas. Essas assimetrias nos circuitos elétricos ou magnéticos causam diferentes correntes em oposição, nas partes de um mesmo enrolamento. Se uma bobina é enrolada de forma concêntrica ao redor do eixo de uma máquina elétrica, então devido aos fluxos de dispersão axiais, presentes na máquina causados por assimetrias elétricas e magnéticas, serão induzidas tensões nesta bobina. Pelo monitoramento da tensão induzida, sua análise espectral pode ser usada para identificar várias assimetrias e condições faltosas. Pela aplicação desta técnica simples e não invasiva é possível detectar assimetrias na gaiola rotórica de uma máquina de indução, como barras quebradas, curtos-circuitos entre espiras nos enrolamentos estatóricos, curtos-circuitos no enrolamento do rotor bobinado, perdas de fase na alimentação, excentricidades, etc.

(Schoen et al., 1995) utilizam a análise espectral de corrente estatórica monofásica para detectar danos nos rolamentos da máquina de indução pela correlação entre vibração e as frequências presentes nas correntes. Trabalham na correlação entre frequências de vibração e de correntes causadas por faltas incipientes. A vantagem do método proposto é que com o monitoramento da corrente estatórica, não é necessário acesso físico ao motor.

Hsu (1995), sugere um método para monitorar faltas rotóricas ou nos enrolamentos do estator de motores de indução a partir do monitoramento do conjugado de entreferro, o qual, sendo produzido pelo fluxo concatenado e pelas correntes, é sensível a um desequilíbrio criado por defeitos. Tal conjugado é diferente do conjugado medido no eixo. O espectro do conjugado permite identificar componentes na frequência dupla da fundamental para defeitos no estator e na frequência dupla do escorregamento para defeitos no rotor. O método não exige paradas especiais para medições específicas, já que o conjugado de entreferro pode ser

calculado durante funcionamento normal, durante um tempo pequeno. A análise é conduzida sobre: (a) operação normal; (b) falta de fase no estator; (c) falta de fase no rotor.

(Schoen et al., 1995) dão uma importante colaboração na identificação de faltas através da análise espectral. Vários métodos têm sido propostos para a detecção de excentricidades e barras rotóricas quebradas, no entanto, tais métodos geralmente assumem que o conjugado de carga é constante. Os autores afirmam que se o conjugado de carga varia com a posição do rotor, então os harmônicos espectrais da corrente do motor produzidos pela carga vão sobrepor os harmônicos produzidos pelas condições de falta naqueles sistemas que medem a corrente estatórica por apenas uma fase. Mostram que os efeitos de um campo magnético sob condições de falta e uma oscilação de conjugado sobre o espectro de corrente do estator não podem ser separados sem conhecimento da posição espacial da falta e do conjugado de carga em relação ao rotor.

Stone e Kapler (1998) afirmam que, em um universo de 7500 falhas detectadas em motores de indução, 37% do total devem-se a problemas no enrolamento estatórico e que, até recentemente nenhum monitoramento on-line efetivo estava disponível para detectar falhas nos enrolamentos de motores e geradores. Sem este monitoramento *on-line*, a máquina necessitaria ser removida de serviço com uma certa frequência para inspeções e testes off-line, para assegurar a condição do enrolamento.

Estudaram também o mecanismo mais comum que originam faltas no enrolamento estatórico: a operação em altas temperaturas, por longo tempo, que causa envelhecimento do isolamento, fazendo com que as camadas de isolante fiquem compactadas e quebradiças. Este processo conduz à criação de bolsões de ar. Também identificaram que, além deste mecanismo, outros como vibração das bobinas ou espiras e poluição (misturas condutoras com umidade, poeira e/ou óleo) também criam os bolsões de ar. Em máquinas acima de 4kV, foi detectada a presença de descargas parciais que são pequenas faíscas que ocorrem no

isolamento, onde quer que os bolsões de ar existam. Assim a presença de descargas parciais indica que uma falta está em andamento no estator, permitindo que atitudes como: manutenção corretiva, mudança de operação ou re-enrolamento da fase sejam tomadas. Isto evita longos desligamentos não planejados e estende a vida útil da máquina.

Por cerca de 40 anos, fabricantes de motores e geradores realizaram medições periódicas das descargas parciais e detectaram problemas de alta tensão nos processos de falha no isolamento dos enrolamentos estatóricos. O método já vem sendo utilizado, com sucesso por empresas como American Electric Power, Virginia Power e Ontario Hydro.

A detecção das descargas parciais em motores acima de 6 kV é feita a partir de um instrumento portátil chamado TGA-B, o qual é utilizado durante operação normal do motor. Separa os pulsos das descargas parciais dos sinais devidos a ruídos e grava a quantidade, a magnitude e a posição das descargas parciais.

Cash, Habetler e Kliman (1998) apresentam uma nova técnica de detectar falhas no isolamento de máquinas polifásicas de corrente alternada, a partir da soma algébrica de três tensões de fase instantâneas em máquinas conectadas em estrela e com neutro acessível. A sensibilidade do método a frequências indesejáveis é resolvida com a utilização de um filtro de harmônicos em torno da frequência fundamental. A grande vantagem da técnica é ser simples na teoria e na prática, garantindo implementação de baixo custo.

(Riley et al., 1998) propõem e avaliam um método de monitoramento on-line da vibração baseado no relacionamento entre harmônicos de corrente e os de vibração. Estudos teóricos, resultados de simulação e resultados experimentais demonstraram que existe um relacionamento linear entre os harmônicos de corrente e o nível de vibração. A constante de proporcionalidade é obtida pelos valores médios de vibração durante o funcionamento normal da máquina.

2.5.1 – A detecção de faltas, utilizando técnicas de inteligência computacional e o método qualitativo

Muitas bibliografias referem-se à Inteligência Artificial, enquanto outras utilizam o termo Inteligência Computacional. Neste trabalho, consideraremos como pertencentes à Inteligência Artificial, os sistemas baseados em regras. As Redes Neurais, Sistemas Nebulosos e Algoritmos genéticos estarão sendo definidos como pertencentes à Inteligência Computacional. Sistemas referidos como Híbridos estarão englobando mais de uma das características citadas.

A modelagem da máquina de indução é complexa não só pelo uso de equações diferenciais não lineares, mas pela multidão de detalhes inerentes às condições normais e de faltas. Se forem considerados aspectos reais como saturação, histerese, variação dos parâmetros com a temperatura e a saturação e imperfeições construtivas, a modelagem torna-se ainda mais complexa. Uma apreciável melhoria pode ser obtida pelo uso de técnicas de inteligência computacional para gerenciar a aquisição e o processamento de dados, bem como incrementar a efetividade do diagnóstico.

(Chow & Yee, 1990) descrevem várias arquiteturas de redes neurais artificiais para aplicação na detecção on-line de faltas incipientes nas máquinas de indução, baseando em medidas do motor e não em rigorosos modelos matemáticos. Afirmam que, com projeto apropriado, as redes neurais artificiais podem ser usadas satisfatoriamente na detecção de faltas incipientes.

(Chow et al., 1991a) projetaram um detector de faltas incipientes para motores de indução de tamanhos pequeno e médio. Utilizaram uma rede neural artificial como detector de faltas e outra como um filtro. Os resultados do treinamento da rede neural podem ser salvos em formato adequado para a utilização em um software detector de faltas em tempo real. A

rede neural de alta ordem (5 entradas), provê precisão tão boa ou melhor que os detectores baseados em redes neurais convencionais (2 entradas), além disto, a expansão para 5 entradas proporcionou redução no tempo de treinamento. O nível de precisão é maior que 95%, nível satisfatório para aplicações em tempo real.

(Filippetti, Franceschini, & Tassoni, 1995) asseguram que uma apreciável melhoria pode ser obtida utilizando técnicas de inteligência computacional, como sistemas especialistas, para gerenciamento da aquisição de dados e processamento. Afirmam que uma rede neural pode substituir de forma mais eficaz os modelos de faltas em máquinas de indução.

Bernieri, Betta e Liguori (1996) afirmam que a maioria das técnicas propostas na literatura é baseada no uso de modelos de caixa preta do sistema em análise: um modelo nominal (estado correto de funcionamento do sistema), alguns modelos de falta e um modelo real (dinamicamente avaliado durante operação). A comparação entre os modelos real e nominal permite detectar a falta, enquanto a comparação entre os modelos real e de falta permite o diagnóstico da falta. Segundo os autores, as abordagens usadas para construção dos modelos têm sido: (a) simulação de circuito equivalente para construir modelos não paramétricos; (b) sistema de identificação para estimar modelos paramétricos.

A aplicação de modelos, para representar faltas, requer o conhecimento profundo de sistemas com complexas características não-lineares, requerendo grande esforços computacionais.

As técnicas de identificação de modelos de sistemas paramétricos oferecem amplo conjunto de soluções. Todos requerem avaliação preliminar da estrutura do modelo do sistema e são geralmente limitados em relação ao sistema em análise e às modalidades de operação.

Nos últimos anos, pesquisas no campo de Redes Neurais Artificiais sugeriram a intensificação no uso dessa metodologia nos sistemas de diagnóstico.

2.5.2 - A utilização de Redes Neurais no monitoramento das condições das máquinas elétricas

(Chow et al., 1990, 1991a, 1991b), identificaram que a corrente estatórica e a velocidade apresentam indicações da condição dos enrolamentos estatóricos e dos rolamentos. Além de serem acessíveis, podem ser medidos com precisão. Devido a isto, estas foram as variáveis escolhidas para a detecção de faltas incipientes em um motor de indução monofásico com rotor em gaiola de esquilo. Dedicaram-se ao estudo de basicamente dois tipos de faltas: (a) faltas no enrolamento estatórico e (b) deterioração dos rolamentos dos mancais.

Descrevem várias arquiteturas de redes neurais artificiais adequadas para a aplicação na detecção on-line de faltas incipientes. Esta técnica não necessita de nenhuma modelagem matemática do motor. As medidas podem ser tomadas diretamente na própria máquina.

Mostram que redes neurais artificiais eliminam a necessidade de compreender com precisão as dinâmicas dos sistemas, o que é necessário em outras abordagens usadas para estimar os parâmetros da máquina e indicar as condições da máquina.

Chow et al. (1990,1991a,1991b) estabeleceram relações entre parâmetros e grandezas mensuráveis - velocidade do motor (W), a corrente elétrica (I), coeficiente de atrito (Bc) e número de espiras equivalente do estator (Nc) – como forma de identificar as falhas na máquina em estudo. Estabeleceram relações entre a velocidade do motor em regime permanente e o coeficiente de atrito e entre a corrente e o número de espiras equivalente no estator, encontrando o seguinte equacionamento (para maiores detalhes, ver o artigo em referência):

$$g(I, W, Nc, Bc) = 0 \quad (9)$$

No entanto, o desenvolvimento de um modelo matemático pode ser substituído pelo uso de uma rede neural que tenha “aprendido” o mapeamento entre entradas e saídas, sendo capaz de identificar as situações para as quais tenha sido treinada e sinalizando com a identificação do defeito ou condição normal correspondente. A rede, uma vez treinada, representa as não-linearidades do mapeamento desejado.

(Chow et al., 1991a) projetaram um detector de faltas, utilizando uma rede neural artificial de alta ordem e outra rede para funcionar como um filtro de medidas feitas durante estados transitórios, utilizando um algoritmo modificado de aprendizado competitivo. Os dados de treinamento foram obtidos a partir de um programa capaz de simular o desempenho do motor de indução monofásico sob diferentes condições, incluindo indutâncias de dispersão, efeitos da saturação, faltas nos enrolamentos estatóricos e nos rolamentos.

Afirmam que a interpretação das condições das faltas é um conceito nebuloso e que a lógica nebulosa está implicitamente incluída no funcionamento da rede neural. Além disto, consideram que a classificação das faltas é muito subjetiva e dependente de vários fatores, o que justifica ser tratada como problema de lógica nebulosa.

Penman e Yin (1992) apresentam uma abordagem utilizando redes neurais artificiais para identificação de faltas em motores de indução trifásicos, partindo do raciocínio de que qualquer condição de operação anormal gera alterações no fluxo de entreferro e que grandezas que são função do fluxo do entreferro reproduzirão tais anormalidades. Devido a isto, selecionaram como parâmetros a serem medidos, a vibração e a corrente.

Utilizaram um detector de faltas baseado em rede neural artificial multicamada, o algoritmo *backpropagation* e a técnica do *momentum* - ver APÊNDICE “C”. Os dados de treinamento foram obtidos sob vários regimes e diferentes condições de carga. A rede foi treinada para reconhecer situações normais de operação, suprimento desequilibrado e perda mecânica.

(Filippetti, Franceschini e Tassoni, 1993) também utilizam, neste trabalho, uma rede neural artificial multicamada, o algoritmo *backpropagation* e a técnica do *momentum* para reduzir o erro na saída e melhorar a taxa de convergência. O objetivo é fazer a detecção e a localização de faltas rotóricas tipo barras quebradas. Utilizam o mapeamento das componentes espectrais de corrente estatórica nas frequências “f” e “(1-2s)f” em relação ao escorregamento e ao número de barras rotóricas quebradas. Chamam a atenção para o fato de que para que uma rede neural seja adequada ao propósito de diagnóstico, deve-se escolher adequadamente as entradas e o critério de normalização de tais entradas.

(Penman & Yin, 1994) utilizam uma rede neural Perceptron Multicamada (MLP), treinada por um algoritmo de aprendizado *backpropagation*, para aprender e reconhecer padrões, sinais de entrada associados a estados de máquinas. Utilizam também, o método não supervisionado - mapeamento de Kohonen (KFM), apresentando as vantagens deste método sobre o supervisionado, dentre as quais, o fato de não necessitarem de muitas iterações no treinamento, serem mais rápidas, não necessitarem de pares entrada/saída previamente definidos e produzirem classificações sem supervisão. Diferentes defeitos de origem elétrica ou mecânica foram submetidos à rede, juntamente aos dados de funcionamento normal. A rede foi capaz de se auto-organizar, tal que padrões pertencentes a uma mesma classe foram agrupados juntos no espaço de saída, além disto está preparada para receber dados no domínio do tempo e também no domínio da frequência. Os resultados demonstram significativa redução nos custos de aquisição de dados associados com o monitoramento de condições.

(Schoen et al., 1995) apresentam um método on-line para detecção de faltas em um motor de indução. Utiliza um sistema especialista baseado em regras, um filtro de frequências e uma rede neural para determinar mudanças no espectro de corrente produzidas por uma falta incipiente. A rede “aprende” as características específicas do motor durante operação normal. O espectro aprendido pode conter muitos harmônicos devido à carga, correspondentes às

condições normais de operação. O filtro seletivo de harmônicos determina que frequências devem ser monitoradas pela rede neural. Esta, por sua vez detecta alterações na condição normal, por comparação com uma condição anterior da máquina e é criado um novo agrupamento pelo algoritmo da rede neural. Após algum tempo a rede sinaliza uma condição de falta potencial se o novo agrupamento persistir. Dado às características descritas, este novo método não requer informações sobre o motor ou a característica da carga para fazer a identificação de uma situação de falta. O método apresentado não requer nenhuma interpretação especial do sinal de corrente do motor, mesmo em presença da carga e de condições de linha desequilibradas. Uma vez treinada a rede, se for detectado um defeito que cause degradação do sinal de corrente por um tempo, o sistema compara as mudanças com o espectro aprendido, indica a condição de falta e dá o alarme quando o desvio for significativo.

(Filippetti, Franceschini, & Tassoni, 1995) utilizam uma rede neural *feedforward* para diagnósticos on-line de faltas rotóricas. A utilização da rede neural não requer modelos das máquinas faltosas nem das formas como operam. Os dados de treinamento podem ser obtidos através das máquinas sob operação em falta ou através do uso de um simulador. Assim como no artigo dos mesmos autores, de 1993, utilizam, neste, o mapeamento das componentes espectrais de corrente estatórica nas frequências “ f ” e “ $(1-2s)f$ ” em relação ao escorregamento e ao número de barras rotóricas quebradas. Desenvolveram um algoritmo auto-ajustável que varia a componente fundamental em correspondência à transformada rápida de Fourier da frequência de saída e, assim, resolveram os problemas de dispersão espectral que obscurece os componentes das bandas laterais. O problema de incluir limiares que representassem dissimetrias construtivas específicas de cada máquina foi resolvida pela abordagem híbrida em termos de dados de treinamento, obtidos a partir de dados de funcionamento normal e utilizando um simulador de faltas. Os dados obtidos durante o funcionamento normal da

máquina já refletem a influência de tais dissimetrias, eliminando a necessidade de outras providências ou considerações.

2.6 – Conclusões

Foi apresentada, neste capítulo, uma síntese dos trabalhos de pesquisadores de vários centros de pesquisa, quanto às metodologias usadas na concepção de sistemas de diagnóstico de faltas em sistemas físicos. Dentro deste contexto situa-se, com destaque, o uso das redes neurais artificiais, as quais podem, uma vez treinadas para isto, realizar todas as etapas do processo supervisório. As pesquisas indicam que as tendências atuais são de controle da generalização das redes neurais, a utilização de modelos com inspiração biológica e o uso de sistemas neurais híbridos.

Vários autores apresentaram artigos que demonstram que as técnicas de inteligência computacional podem conferir uma apreciável melhoria no gerenciamento, aquisição e processamento de dados, tornando o diagnóstico mais rápido e efetivo.

Poderíamos sintetizar ainda alguns aspectos relativos às redes neurais artificiais:

(a) destacam-se pela rapidez na implementação, altas velocidades operacionais e pelo fato de não requererem nenhuma modelagem matemática do processo, mas treinamento e validação com dados experimentais; (b) aprendem através de exemplos; (c) com o uso de redes neurais artificiais evita-se a necessidade de compreender com precisão a dinâmica dos sistemas, o que é interessante no caso de sistemas como o das máquinas elétricas, cuja modelagem apresenta complexidades inerentes à resolução de equações diferenciais e à consideração dos fenômenos vinculados a seu funcionamento.

CAPÍTULO 3

DESENVOLVIMENTO DO PROTÓTIPO E MONTAGEM DA BASE DE DADOS

3.1 - Introdução

Este capítulo é dedicado aos procedimentos metodológicos referentes ao tratamento experimental, à obtenção de dados, pressupostos e limitações considerados no desenvolvimento do simulador real de faltas elétricas, obtido a partir de adaptações feitas nos enrolamentos do estator de um motor de indução trifásico. Os resultados obtidos a partir das simulações reais feitas serão usados para treinar redes neurais artificiais.

3.2 - O simulador físico de faltas

A bibliografia estudada no capítulo 2, em geral, aborda o estudo de faltas em motores de indução monofásicos, no entanto, este trabalho destina-se ao estudo de faltas em motores trifásicos. Como apresentado no item 2.5, vários autores fizeram estudos considerando a corrente estatórica como fonte de informações sobre o estado das máquinas. Apesar do fluxo magnético fornecer informações mais completas sobre a condição de funcionamento dos

motores, a obtenção da corrente é feita de forma mais acessível, não necessitando de técnicas invasivas para sua obtenção.

Objetivando a detecção de faltas incipientes e a manutenção preventiva, diferentes métodos invasivos e não-invasivos podem ser utilizados. Os métodos não-invasivos são os preferidos por não requerem desmontagem do motor ou alterações no projeto.

Krause (1986) demonstra que grandezas como o fluxo concatenado, conjugado e corrente elétrica são funções do número de espiras equivalentes em cada fase, porque parâmetros como resistências e indutâncias também o são.

Tendo em vista este conhecimento, as limitações que simplificações e aproximações assumidas, no desenvolvimento do modelo matemático de uma máquina de indução impõem e, considerando ainda que a utilização de dados de medições reais confere maior confiabilidade ao projeto, decidiu-se pela utilização de dados de simulação física de curtos-circuitos, utilizando para este fim uma máquina de indução trifásica, rotor bobinado, adaptada para a realização de ensaios de curto.

As faltas aqui estudadas são, geralmente, atribuídas à deterioração do isolamento, provenientes de aquecimento por sobrecargas, descargas disruptivas entre espiras provocada pelo envelhecimento natural do material isolante ou por sobretensão aplicada às fases da máquina.

Os ensaios foram feitos utilizando-se um motor de indução trifásico com rotor bobinado, cedido pelo Departamento de Engenharia Elétrica da PUC-MG (FIG. 2). A ficha técnica deste motor é apresentada na TAB. 2. As adaptações feitas consistiram em re-enrolar o bobinado do estator, disponibilizando externamente terminais de cada uma das bobinas que compõem o enrolamento (FIG. 3 e 4). Foi também confeccionado um rotor em gaiola para utilização posterior junto com este estator, visando a continuidade das pesquisas. No

enrolamento do estator foram utilizadas bobinas de 7 e 14 espiras (TAB. 3). As TAB. 4 e 5 apresentam os dados eletromecânicos (elétricos e geométricos) de projeto do motor.

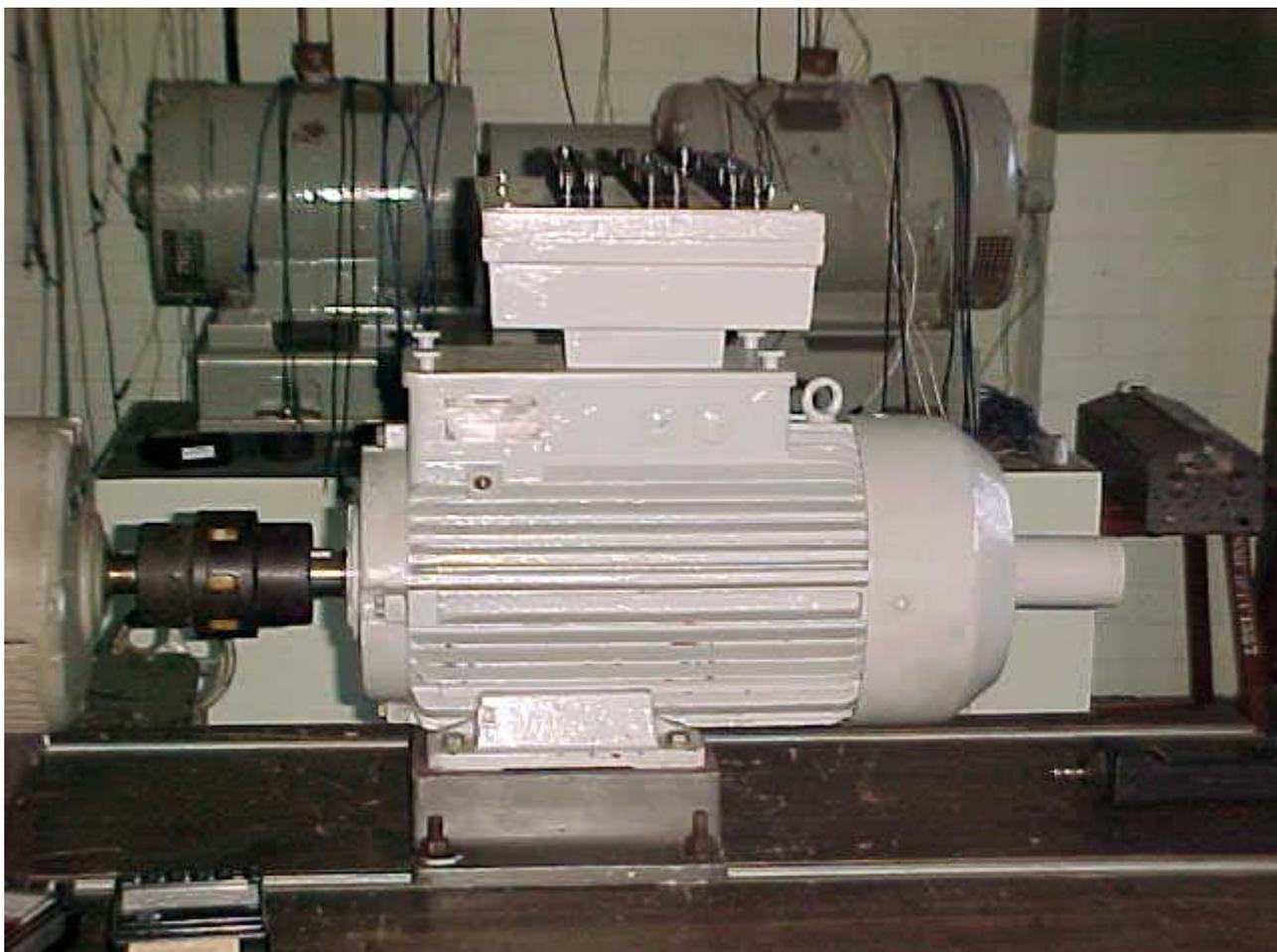


FIGURA 2 - Motor de indução trifásico, com rotor bobinado, utilizado como simulador físico de faltas, neste trabalho.

TABELA 2

Ficha técnica do motor utilizado para obtenção dos dados experimentais

Dados da placa do motor	Características
Marca	SIEMENS
Tipo	1 LS 3133 – 4 AA 90 - Z
Modelo	132 M
Potência	4.6 kW
Rotação	1700 rpm
Tensões estatóricas	127/220 V

Corrente estatórica	34,3/19,8 A
Tensão rotórica	150 V - Y
Corrente rotórica	21 A
Frequência	60 HZ
Fator de potência	0,80
Forma construtiva	B3
Índice de proteção	IP 44

TABELA 3

Conformação das bobinas do estator do motor de indução utilizado na pesquisa.

NÚMERO DE ESPIRAS	BOBINAS
14	1-2; 3-4; 9-10; 11-12
7	5-6; 7-8; 13-14; 15-16

TABELA 4

Dados eletromecânicos (elétricos e geométricos) do motor utilizado para obtenção dos dados experimentais desta dissertação

DADOS ELETROMECAÑICOS (ELÉTRICOS)	ESTATOR	ROTOR BOBINADO
Número de grupos	12	28
No. de bobinas por grupo	2 e 1	1
Passo Bobinas	1: 8: 10 e 1: 8	1
No. de espiras por ranhura	14	1
Tipo bobinado	Concêntrico	Imbricado
No. camadas	1	1
Bitola condutor	4 x 19 AWG	1/8" x 5/8" x 181 mm



FIGURA 3 – Adaptações feitas no motor mostrado na FIGURA 2, para disponibilizar externamente os terminais de cada uma das bobinas que compõem o enrolamento.

TABELA 5

Dados eletromecânicos (geométricos) do motor utilizado como protótipo nesta dissertação.

DADOS ELETROMECAÑICOS GEOMÉTRICOS	ESTATOR	ROTOR BOBINADO
Número de ranhuras	36	48
Diâmetro externo	200 mm	128,2 mm
Diâmetro interno	129,5 mm	50 mm
Comprimento axial	101 mm	100 mm
Abertura ranhura	3,5 mm	
Altura média coroa	19 mm	
Largura média dente	6 mm	



FIGURA 4 – Detalhe da Figura 3, mostrando terminais externos das bobinas 7-8 até 15-16 do estator e os terminais rotóricos.

3.3 - Tratamento Experimental

Foram realizados os ensaios convencionais de rotor bloqueado (TAB. 6) e a vazio (TAB. 7). As medições do ensaio a vazio foram feitas em 220 V. O APÊNDICE B apresenta os parâmetros e as perdas júllicas desta máquina de indução, calculados a partir desses ensaios, da mesma forma que em um transformador.

TABELA 6
Dados medidos durante o ensaio de rotor bloqueado.

Tensão média entre Fases (V)	Corrente na fase "A" (A)	Corrente na fase "B" (A)	Corrente na fase "C" (A)	Potência (W)
55	19.6	19.2	19.9	800

A TAB. 7 também apresenta outros valores obtidos em outro ensaio feito sob frequência constante e tensões variáveis, objetivando obter referenciais do funcionamento do motor a vazio, sob condições normais de funcionamento, quando submetido a tensões inferiores às suas tensões nominais.

TABELA 7
Dados obtidos a partir do ensaio a vazio, com rotor aberto, frequência constante e tensão variável.

Tensão de Linha (V)	Corrente na fase "A" (A)	Corrente na fase "B" (A)	Corrente na fase "C" (A)	Potência (W)	Corrente média (A)
220	10.3	9.8	10	100	10,0
210	8.8	8.6	8.8	90	8,73
200	7.7	7.5	7.7	80	7,63
190	6.8	6.7	6.8	70	6,77
180	6.2	5.9	6.1	56	6,05
170	5.5	5.3	5.5	55	5,43
160	5	4.8	5	43	4,93
150	4.6	4.4	4.6	40	4,53
140	4.2	4.1	4.2	38	4,17
130	3.9	3.8	3.9	30	3,87

3.3.1 Simulação de curtos-circuitos nos enrolamentos do estator

Os enrolamentos do estator da máquina utilizada nos ensaios são compostos por conjuntos de bobinas, com diferentes números de espiras (TAB. 3). As adaptações feitas conforme FIG. 3 e 4 permitiram a aplicação de curtos-circuitos por grupos de espiras no enrolamento do estator, simulando diferentes curtos-circuitos. Os ensaios realizados não foram destrutivos, pois em todos, a tensão aplicada foi aumentada até que quaisquer correntes circulantes nas fases ou nas bobinas atingissem o limite de 20 Ampères que é o valor da corrente nominal estatórica. Em cada ensaio, foram anotados, sistematicamente, os valores das tensões e correntes correspondentes a, no mínimo, 6 diferentes valores de tensões. Provendo alimentação sob frequência nominal e tensões variáveis, a partir de 0 Volts, foram realizados os ensaios apresentados na TAB. 8.

TABELA 8
Ensaio de curto-circuito realizados no motor protótipo

Número de ordem da Classe	Número de espiras em curto-circuito por fase		
	A	B	C
1	0	0	0
2	0	0	28
3	0	0	35
4	0	0	42
5	0	0	70
6	0	0	77
7	0	7	0
8	0	7	7
9	0	7	14
10	0	14	0
11	0	14	7
12	0	14	14
13	0	14	35
14	0	28	0
15	0	28	14

16	0	35	0
17	7	0	0
18	7	7	0
19	7	7	7
20	7	7	14
21	7	14	0
22	7	14	7
23	7	14	14
24	14	0	0
25	14	7	0
26	14	7	7
27	14	7	14
28	14	14	0
29	14	14	14
30	14	28	0
31	28	0	0
32	28	14	0
33	35	0	0
34	35	0	14
35	35	14	0
36	35	28	0
37	42	0	0
38	56	0	0
39	70	0	0
40	77	0	0

Foram realizados curtos-circuitos envolvendo diferentes números de espiras por fase. As FIG. 5, 6 e 7 ilustram como foram utilizados os amperímetros e feitas as ligações para as simulações de: (a) curtos-circuitos de 1 ou mais bobinas, em uma única fase; (b) curtos-circuitos aplicados a duas fases simultaneamente, porém de forma independente; (c) idem para curtos-circuitos envolvendo 1 ou mais bobinas em três fases simultaneamente. Os dados obtidos através dos ensaios foram organizados em um banco de dados, contendo as tensões de linha, as correntes por fase “Y” e o número de espiras em curto por fase. Este conjunto de dados subsidiou o processo de treinamento das redes neurais artificiais, conforme capítulo 4.

Procedimentos adotados para a obtenção dos cerca de 1200 ensaios: (a) foram registrados os valores médios das tensões de linha e das correntes de linha das 3 fases, tendo sido monitoradas, com amperímetros, o nível das correntes internas às bobinas em curto; (b) durante a realização de cada ensaio, variou-se o número de espiras em curto-circuito no

enrolamento do estator; (c) todos os ensaios foram realizados, estando a máquina sem carga, objetivando minimizar efeitos externos originários em conexões com cargas, como desalinhamento e desbalanceamento.

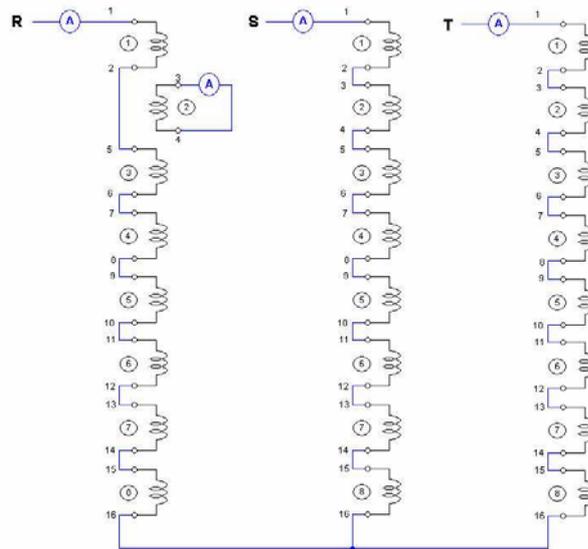


FIGURA 5 - Esquema de ligação para simulação de falta em uma bobina da fase “A”.

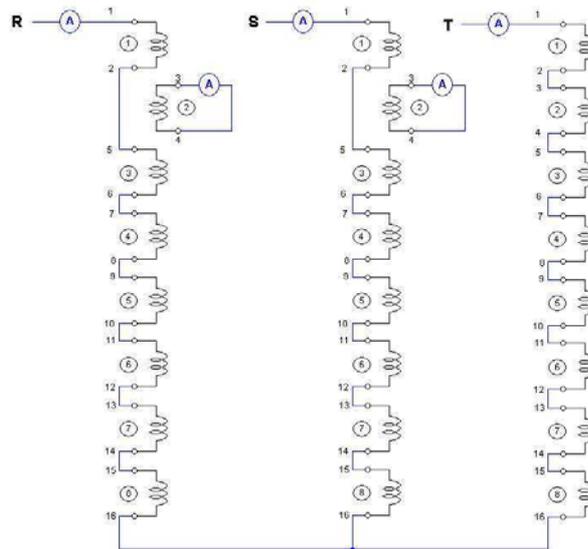


FIGURA 6 - Esquema de ligação para simulação de falta em duas bobinas, em duas fases diferentes, porém, ao mesmo tempo.

Conforme visto no capítulo 2, o levantamento do estado-da-arte apontou que, nos últimos anos, a inteligência computacional tem sido muito utilizada no diagnóstico de problemas, com destaque para a utilização de redes neurais artificiais em tarefas como: classificação de padrões, aproximação de funções, análise de dados, memorização por associações, compressão de dados, predição.

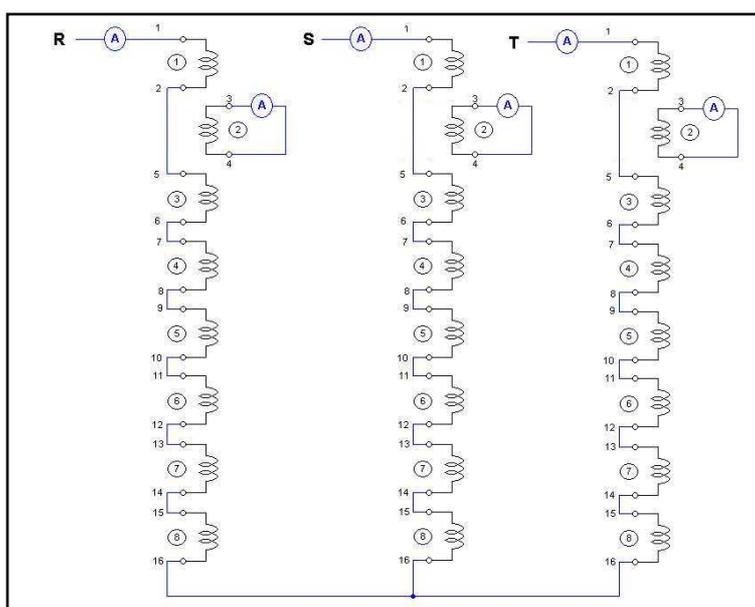


FIGURA 7 - Esquema de ligação do estator para simular falta em três bobinas, uma por fase, ao mesmo tempo.

O estudo de faltas em motores de indução trifásicos pode ser entendido como um problema de classificação de padrões, já que os padrões de comportamento dos motores frente às faltas podem ser submetidos às redes neurais para treinamento. Esta é uma aplicação das redes neurais em Engenharia, no entanto, há uma ampla variedade de outras aplicações em: Computação, Matemática, Física, Psicologia, Biologia, Linguística, Filosofia, Medicina, dentre outras. Como exemplo, podemos citar Chow & Yee (1990,1991a) que comentam sobre a utilização de redes neurais na detecção de faltas, diagnósticos médicos, modelagem de sistemas dinâmicos, controle de robôs, dentre outras.

Alguns aspectos que favorecem o uso de redes neurais são: as inéditas capacidades de aprendizado e de generalização, a adaptabilidade, a tolerância a falhas, a simplicidade e rapidez de implementação, a alta velocidade operacional, bem como o fato de dispensarem a modelagem precisa do funcionamento dinâmico dos sistemas.

Conforme será visto no capítulo 4, a base de dados, formatada inicialmente conforme leiaute apresentado na TAB. 9, foi adaptada para aplicação em diferentes arquiteturas de redes neurais artificiais, o que permitiu vários estudos, inclusive de desempenho das referidas redes aplicadas ao presente problema.

TABELA 9

Leiaute básico da base de dados gerada a partir dos ensaios realizados no motor protótipo.

Ordem	Identificador	Significado
1 ^a	Vab	Valor eficaz da tensão de linha entre as fases.
2 ^a	Ia	Valor eficaz da Corrente de linha na fase "A".
3 ^a	Ib	Valor eficaz da Corrente de linha na fase "B".
4 ^a	Ic	Valor eficaz da Corrente de linha na fase "C".
5 ^a	Na	N ^o . de espiras em curto na fase "A".
6 ^a	Nb	N ^o . de espiras em curto na fase "B".
7 ^a	Nc	N ^o . de espiras em curto na fase "C".

CAPÍTULO 4

MÉTODO QUALITATIVO DE DETECÇÃO DE FALTAS EM MOTORES DE INDUÇÃO, UTILIZANDO REDES NEURAIAS

4.1 – Introdução

Conforme mencionado no capítulo 1 e abordado no capítulo 2, optou-se pela utilização do modelo qualitativo, para diagnosticar as condições de normalidade e anormalidade no funcionamento de um motor de indução, objeto deste estudo. Com base na vasta bibliografia consultada e sintetizada no capítulo 2, concluiu-se que, em se tratando de um problema complexo, uma alternativa interessante seria utilizar inteligência computacional, com ênfase em redes neurais, para abordar tal modelo.

No APÊNDICE C, o leitor poderá ter acesso a uma revisão bibliográfica sobre as redes neurais estudadas.

4.2 – Considerações Gerais

No presente capítulo, as condições gerais, as particulares, os termos de uso específico dentro do contexto e os esclarecimentos prévios e complementares serão apresentados, objetivando dar maiores subsídios para o entendimento das simulações, resultados e análises apresentadas nas seções 4.10 a 4.13.

A partir dos dados obtidos conforme simulações físicas descritas no capítulo 3, foi montada a base de dados, faltava definir que tipo e topologia de rede neural melhor resultado apresentaria para o problema de fornecer um diagnóstico das condições normais e anormais de funcionamento do motor.

4.3 – Escolhendo os tipos de redes neurais

Qual seria a configuração e o tipo de rede neural apropriados para uma representação do processo físico de ocorrência de faltas elétricas em motores de indução?

A literatura consultada sugeria o estudo de diferentes tipos de redes neurais artificiais, explorando-se, neste trabalho, 3 diferentes paradigmas de aprendizagem e 4 arquiteturas, conforme demonstrado na TAB. 10, objetivando conhecer os desempenhos das redes tipo *feedforward* quando comparados àqueles das redes tipo memória associativa.

Devido a isto foram escolhidas duas redes *feedforward*: a perceptron multicamada e a RBF, bem como duas recorrentes: a Memória Associativa Bidirecional (BAM) e a Hopfield.

TABELA 10

Diferentes redes neurais estudadas em suas aplicações ao problema de faltas em motores de indução trifásicos.

TIPO DE APRENDIZADO	REGRA BÁSICA DE APRENDIZAGEM	ARQUITETURA	ALGORITMO DE APRENDIZAGEM
Supervisionado	Regra Delta (Treinamento com Mínimos quadrados)	<i>Feedforward</i> Multi-camada	<i>Backpropagation</i>
Não supervisionado	Hebb Treinamento estático	Rede Hopfield Rede BAM	Aprendizagem de Memória Associativa Aprendizagem de Memória Associativa Bidirecional
Híbrido	1 ^a . Etapa: não supervisionado; 2 ^a . Etapa: supervisionado.	Rede Função de Base radial (caso especial de <i>Feedforward</i> Multicamada).	Algoritmo de aprendizado especial da Rede Função de Base Radial

Como apresentado no APÊNDICE C, nem todas as redes neurais executam o conceito tradicional de treinamento. Redes neurais “aprendem” de diferentes formas, bem como apresentam formas diferenciadas de recuperar os dados “aprendidos”. Assim, podemos destacar as seguintes metodologias de treinamento:

1. Treinamento estático – o ajuste de pesos ocorre seguindo uma regra determinada e independe de algum critério externo;

2. Treinamento dinâmico – o ajuste de pesos ocorre objetivando a diminuição do erro entre a saída desejada e a obtida. O processo de treinamento continua enquanto tal erro não alcançar um objetivo pré-determinado que pode ser o número máximo de iterações, o tempo de processamento, o processo de convergência, dentre outros;
3. Treinamento híbrido – o ajuste de pesos é feito em duas etapas, sendo uma delas definidas por regras (não-supervisionada) e outra feita objetivando a minimização do erro (supervisionada).

Também devem ser mencionadas as diferentes metodologias de recuperação de padrões, mediante a apresentação de vetores de entrada e dos pesos ajustados durante a etapa de treinamento:

1. Avaliação estática – a recuperação de uma saída é feita de maneira direta, a partir de simples operações matriciais;
2. Avaliação dinâmica – a recuperação de uma saída é feita de maneira iterativa, ou seja, são realizadas operações de modo a encontrar o padrão armazenado que mais se aproxime daquele representado pelo vetor de entrada.

Exemplifiquemos os conceitos, utilizando a Rede Perceptron Multicamada *feedforward*, treinada com algoritmo *backpropagation*. Esta é uma rede classificada como não-recorrente. Não é uma memória associativa como a Hopfield e a BAM. Durante a fase de treinamento, gera uma matriz de pesos. Uma vez treinada, ao receber outros padrões, na fase de recuperação, a rede gera os valores de saída, função dos pesos e das entradas. Assim, cada entrada gera sempre a mesma saída; a fase de aprendizado pode ser classificada como um processo iterativo, enquanto a fase de recuperação da rede pode ser considerada um processo estático. O processo de treinamento é classificado como supervisionado.

4.4 – Pré-processamento da massa de dados

Este é um dos mecanismos utilizados para melhorar a velocidade do treinamento e reduzir as dimensões do espaço de entrada.

Neste trabalho, o pré-processamento dos vetores de entrada é diferenciado para cada tipo de rede, conforme será visto nas próximas seções deste capítulo. Os pré-processamentos comuns a todas as redes simuladas foram:

- Utilização de valores “pu”: para aumentar a aplicação da presente pesquisa, os dados dos vetores referentes às correntes e tensões foram normalizados em função dos valores nominais do motor tomado como protótipo;
- Para aumentar a quantidade de dados a vazio, sem falta, os dados obtidos sob esta condição foram tabulados e colocados na forma de gráfico (FIG. 8). Aplicando a interpolação por um polinômio, foi possível obter os valores das correntes para outros valores de tensão além daqueles medidos durante as simulações físicas, a vazio, no protótipo real.

Como parte do pré-processamento, o parâmetro I_N – corrente que circula em condições normais de funcionamento, sem carga e sob tensão V – foi obtido a partir de um ensaio a vazio. Os valores medidos geraram uma curva que foi interpolada por um polinômio que gerou a coluna de corrente nominal para quaisquer tensões aplicadas nos ensaios em curto-circuito.

Abaixo segue o polinômio (FIG. 8) resultante da análise da curva:

$$\boxed{y = 0.00000248416480x^3 - 0.00070177657000x^2 + 0.07931704500000x - 0.00104040170000} \quad (10)$$

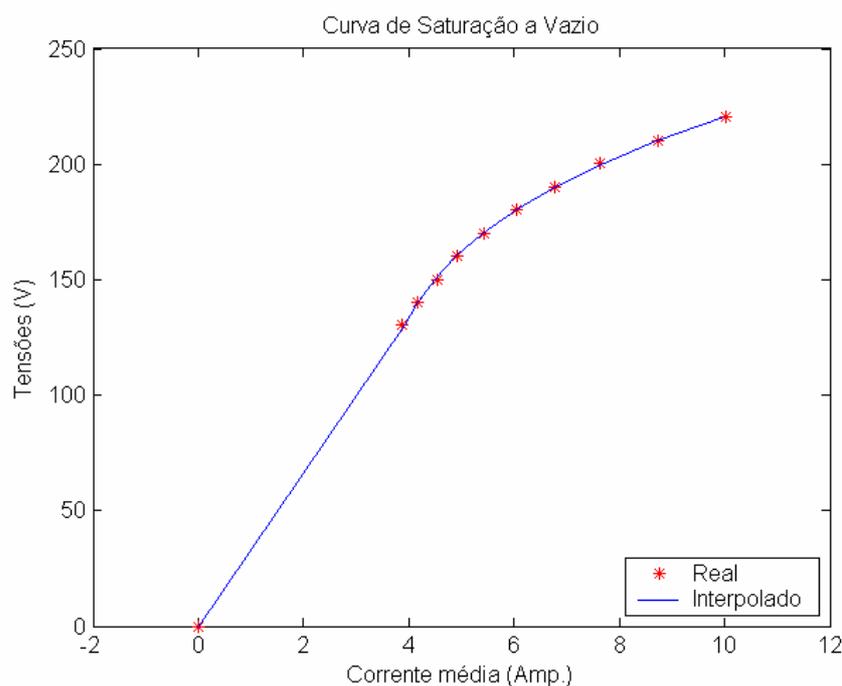


FIGURA 8 – Curva obtida a partir da interpolação da curva tensão x corrente a vazio.

4.5 – Obtenção dos conjuntos de treinamento e validação

Chow (1997) recomenda que, para treinar uma rede neural, seja utilizado um conjunto de dados estatisticamente significativo, que deve cobrir toda a faixa de condições operacionais, incluindo suas variações dinâmicas. A quantidade de dados deve ser grande o suficiente para representar bem o problema em estudo. As variáveis escolhidas devem ser acessíveis, confiáveis e sensíveis para que se tenha uma adequada modelagem da rede. Além disto, conjuntos de dados devem ser separados para fazer o treinamento e a validação (teste) da rede.

Tradicionalmente, do conjunto total de dados, um percentual variando entre 50 e 90% é utilizado para treinamento. Tais dados devem ser escolhidos aleatoriamente, para que a rede

“aprenda” as regras e não, simplesmente, “decore” exemplos, embora, nem todas as redes tenham capacidade de fazer a generalização, funcionando, realmente, como memória. Conseqüentemente, um percentual entre 50 e 10% do conjunto inicial de dados deve ser separado para fazer a validação da rede, permitindo explorar diferentes configurações, quanto à capacidade de generalizar.

É necessário também, fazer análises do funcionamento da rede neural, a partir dos chamados erros aparente e verdadeiro. Podem ser definidos como erros aparentes os erros obtidos sobre o conjunto de treinamento, que são otimizados pelo uso de um algoritmo de treinamento. Já o erro verdadeiro é obtido na chamada fase de validação, quando se submete à rede um conjunto de dados independentes daqueles utilizados durante o treinamento. A análise dos erros obtidos na fase de validação permite comparar o desempenho de diferentes redes.

Neste trabalho, a massa de dados total recebeu o pré-processamento adequado às características próprias das simulações realizadas. Em seguida, em todos os casos, houve separação aleatória dos dados em dois conjuntos, utilizando o comando “randperm” do MATLAB: 75% dos dados foram usados para treinamento e 25% para validação.

4.6 – Procedimentos gerais para melhoria das simulações das redes neurais artificiais

A implementação de cada um dos diferentes tipos de redes neurais não foi feita em única tentativa, pois, não se obtém, independente das características, na primeira implementação, uma rede neural com bom desempenho em termos de convergência, velocidade e capacidade de generalização. Quando uma rede, após a fase de treinamento,

apresenta baixa taxa de acertos, conclui-se que tal configuração de rede é incapaz de realizar a tarefa que se estabeleceu como objetivo. Pode-se adotar uma das alternativas: (a) alterar a estrutura e os parâmetros até encontrar uma estrutura capaz de atender ao objetivo estabelecido, inclusive com adequação do tempo requerido para treinamento; (b) mudar o algoritmo de treinamento.

Quando o conjunto de dados de treinamento possui muito ruído ou os atributos dos padrões de entrada não possuem informações suficientes, pode-se: (a) aumentar a quantidade de dados, (b) aumentar as informações contidas nos vetores de entrada ou (c) mudar os objetivos pretendidos.

Nas simulações feitas, as diferentes alternativas citadas foram utilizadas, objetivando a melhoria do desempenho das redes.

O ambiente MATLAB da MathWorks versão 6.5, *toolbox* de Redes Neurais, é amigável e bastante favorável à realização de pesquisa, baseada em simulação, por possuir funções desenvolvidas especialmente para manipulação de matrizes e para simulação das diferentes redes neurais artificiais. Disponibiliza 3 parâmetros distintos para acompanhamento e interrupção das simulações: (a) número máximo de iterações ou épocas (*Epoch*); (b) valor do erro médio quadrático admissível (MSE) e (c) a norma do vetor gradiente (*Gradient*) definida também como objetivo.

Assim, o treinamento é interrompido:

- Se o número máximo de iterações for alcançado;
 - Se o valor do erro médio quadrático alcançar o valor estabelecido como objetivo;
 - Se a norma do vetor gradiente ficar menor que o valor especificado, ou seja, se a norma assumir valor infinitesimal, então teremos encontrado um mínimo, que deverá ser avaliado porque estará indicando a ocorrência de um mínimo local ou global.
-

- a) O *early-stopping* é um importante recurso que o MATLAB oferece, conhecido como parada antecipada, conforme tradução de Haykin (2001). O uso deste recurso permite parar o treinamento quando o erro de validação começa a aumentar, além de melhorar o desempenho da rede em relação à capacidade de generalizar, a qual pressupõe o aprendizado de uma “regra” quando lhe são apresentados padrões que não estavam presentes no conjunto de treinamento.

O MATLAB 6.5 permite, a partir de comandos padrões, utilizar a parada antecipada conforme exemplo de código fonte da TAB.11.

TABELA 11

Código que demonstra a utilização do recurso do MATLAB 6.5, denominado parada antecipada

```
% Vetores de Treinamento
xtrain = [-1:0.05:1];
ytrain = sin(2*pi*xtrain) + 0.1*randn(size(xtrain));

% Vetores de Validação
val.P = [-0.5:.01:0.5];
val.T = sin(2*pi*val.P) + 0.1*randn(size(val.P));

% Inicialização estrutural da rede
% NEWFF(PR,[S1 S2...SN1},{TF1 TF2...TFN1},BTF) recebe:
% PR - matriz Rx2 de valores mínimos e máximos para as R entradas
% Si - Tamanho da i-ésima camada, para N1 camadas
% TFi - Função de ativação da i-ésima camada, default = 'tansig'
% BTF - Função de treinamento por retropropagação, default = 'trainlm'
%
% PR = [-1 1], ou seja, as entradas variam nesse intervalo;
% Si = [20 1], ou seja, 20 nós na camada escondida e 1 na camada de saída;
% TFi = {'logsig','purelin'}, ou seja, a função de ativação da camada
% intermediária é a sigmóide('logsig') e a da saída é linear('purelin');
% BTF = 'trainlm', ou seja, algoritmo Levenberg-Marquardt backpropagation
net = newff([-1 1],[20 1],{'logsig','purelin'},'trainlm');

% Número máximo de épocas
net.trainParam.epochs = 100;

% Inicia Treinamento com validação cruzada
% TRAIN(NET,P,T,Pi,Ai,VV) recebe,
% NET - Estrutura do tipo rede que foi criada com NEWFF.
% P - Matriz de treinamento contendo a(s) entrada(s).
% T - Matriz de treinamento contendo a(s) saída(s) desejada(s)
```

```
% Pi - Condições iniciais de atraso na(s) entrada(s), default = zeros.  
% Ai - Condições iniciais de atraso na(s) camada(s), default = zeros.  
% VV - Estrutura contendo as matrizes de validação  
net = train(net,xtrain,ytrain,[],[],val);  
  
plot(xval,sim(net,xval))
```

Para melhor entendimento, considere-se que:

- b) No treinamento convencional, o erro médio quadrado começa grande e decresce com o aumento do número de iterações até encontrar um mínimo na superfície de erro. Como visto previamente, devem ser definidos parâmetros limites, os quais determinam o momento de parar o treinamento. O treinamento é, assim, monitorado para evitar a perda de tempo na manutenção de um treinamento que tenda a divergir, ao invés de convergir. Além disto, o teste com a massa de validação permite ir conhecendo a capacidade de generalizar, além de evitar que a rede acabe excessivamente identificada com os dados de treinamento, caracterizando o chamado *overfitting*.

Os subconjuntos de treinamento e de validação precisam ser organizados em duas matrizes: uma contendo os padrões de entrada e outra contendo os padrões de saída.

O processo de parada antecipada consiste em treinar a rede, aplicando o teste com o subconjunto de validação, de forma periódica. Assim, após uma estimação do erro de treinamento, segue-se uma etapa de aplicação do subconjunto de validação, na qual a rede opera, permitindo a obtenção do erro de validação.

A utilização do *early-stopping* permite que, em um mesmo gráfico sejam plotados os erros médios quadrados do treinamento e da validação, como função do número de iterações. A análise das curvas obtidas permite definir onde o erro obtido com o subconjunto de validação começa a crescer, opondo-se ao erro obtido com o conjunto de treinamento que continua decaindo. Este ponto é o considerado sensato para encerrar o treinamento, considerando-se que o que a rede aprende, após este ponto é essencialmente vinculado ao

ruído. Assim, quando o erro de validação divergir então o treinamento será abortado independente do fato do erro de treinamento estar sendo minimizado.

Na FIG.9 pode ser visualizado como termina o treinamento através do *early-stopping*. Existe uma deficiência nesse tipo de treinamento que é o fato de não ser garantido que o primeiro mínimo do erro de validação seja o mínimo global de validação.

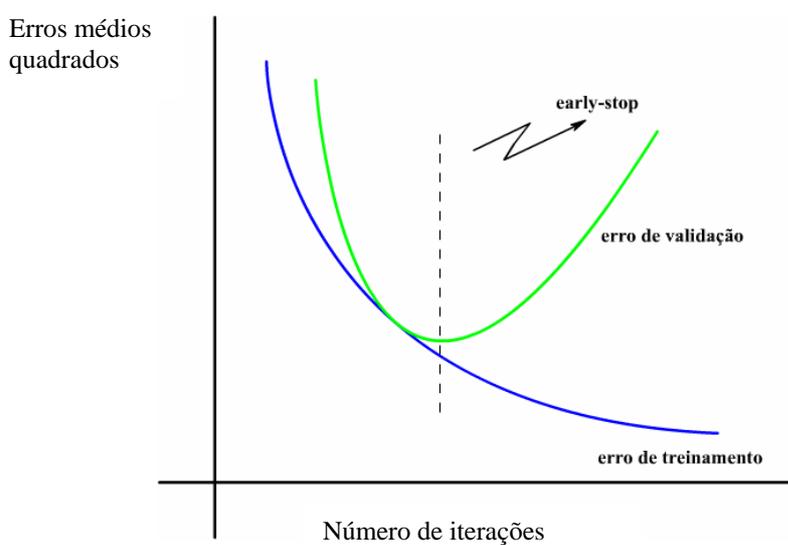


FIGURA 9 – Treinamento com parada antecipada

4.7 – Capacidade de Generalização

Uma das mais importantes características de uma rede neural diz respeito à sua capacidade de generalizar. Esta é uma habilidade que as redes desenvolvem em função de suas estruturas e treinamento.

Shapiro (2003) define a capacidade de generalizar como sendo uma habilidade que a rede desenvolve de usar uma regra para classificar padrões que estejam diferentes daqueles com os quais a rede foi treinada.

Uma rede pode funcionar perfeitamente como memória e não ser capaz de generalizar.

De mesma forma, uma rede pode ter erro bem pequeno e também não ser capaz de generalizar.

Modelos muito simples não conseguem atender à necessidade, muito complexos não conseguem aprender a partir de pequenas quantidades de dados. É necessário selecionar a adequada complexidade fazendo a “regularização”, que paralisa o treinamento quando o erro de validação começa a crescer.

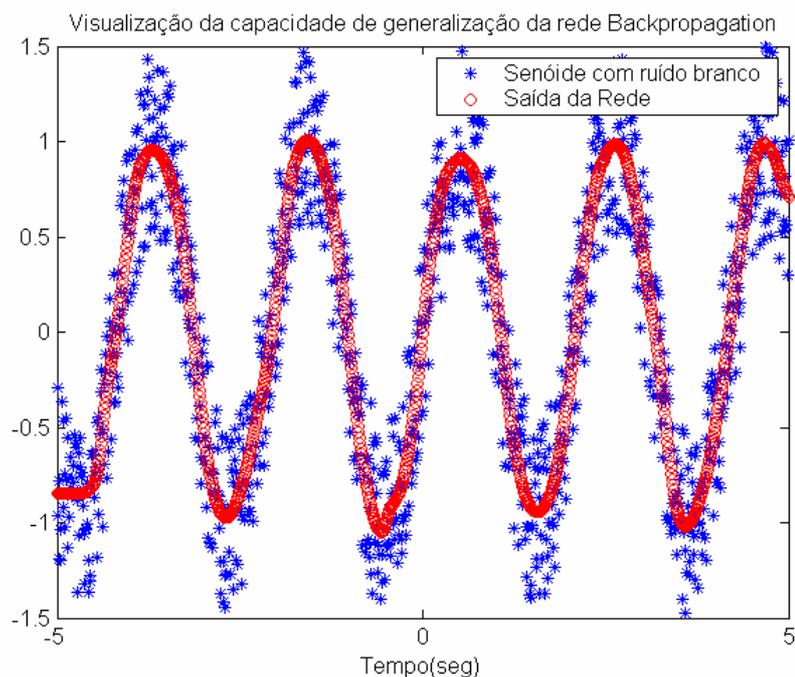


FIGURA 10 – Visualização de tela de saída de programa feito em MATLAB, demonstrando, através da curva senoidal, como uma rede faz a generalização, utilizando a Perceptron Multicamada.

Haykin (2001, p. 232) diz que uma rede “generaliza bem quando o mapeamento de entrada-saída computado pela rede for correto (...) para dados de teste não utilizados para a

criação ou treinamento da rede”. Segundo o mesmo autor a generalização é influenciada pelo tamanho e representatividade do conjunto de treinamento; a arquitetura da RNA e a complexidade do problema a se resolvido.

A FIG.10 exemplifica o conceito de generalizar: ser capaz de identificar ou classificar padrões de entrada distorcidos por ruídos ou incompletos, aplicando a regra implícita com a qual foi treinada.

É desejável que uma rede neural tenha capacidade de generalizar, porém, tal capacidade é fortemente influenciada pela arquitetura da rede, sua complexidade, o tamanho do conjunto de treinamento e a forma como os dados foram divididos.

4.8 – Parâmetros para avaliação das redes

Estabelecemos os parâmetros para avaliação das redes implementadas:

- a) Tempo de treinamento ou recuperação de informações;
- b) Tempo de validação;
- c) Facilidade de implementação;
- d) Estatística de erros e acertos quando submetida à massa de validação;
- e) Capacidade de realizar a generalização, frente à ocorrência de padrões diferentes daqueles contidos na massa de treinamento. Utilizou-se para este fim, a massa de dados de validação.

O uso de histogramas de erro objetivou favorecer a visualização da distribuição do erro entre a saída real e a desejada. O acerto percentual só foi avaliado quando seu uso era coerente.

4.9 – Simulações e resultados obtidos

As simulações, gráficos e tabelas apresentados nesta seção foram obtidos, utilizando-se os recursos do MATLAB 6.5 Release 13, *toolbox* de Redes Neurais Artificiais.

Devido à qualidade dos dados, as massas de dados necessitaram de adaptações, conforme a natureza da modelagem e as características específicas de cada rede neural simulada.

As primeiras simulações foram feitas, permitindo a localização do defeito, bem como, informando o número de espiras em curto, em relação a cada uma das fases.

A base de dados foi obtida em laboratório, a partir das simulações realizadas no motor de indução tomado como simulador físico de defeitos – ver Capítulo 3. Foram utilizados amperímetros e voltímetros. Os dados obtidos não apresentaram intervalos regulares, a homogeneidade e coerência de dados obtidos através do uso de placas de aquisição de dados. A consequência foi obtermos uma modesta quantidade de dados, com características específicas, para treinamento e validação de redes neurais. Assim, em função da base de dados disponível e para obter uma modelagem mais simples e eficiente, foi necessário desprezar, como saída das redes, a informação sobre o número de espiras em curto.

Nesta seção são apresentadas as simulações feitas, utilizando-se diferentes tipos de redes e topologias. O objetivo é encontrar um modelo de rede que melhor desempenho apresente quando submetido à base de dados disponível.

A TAB.12 apresenta uma síntese das experiências feitas, sendo que o termo “simulação” é aplicado para referirmo-nos às diferentes topologias ou aos diferentes formatos de apresentação dos dados dos padrões de entrada ou de saída.

Os termos técnicos aqui utilizados foram mencionados no início deste capítulo.

TABELA 12
Síntese das simulações realizadas

IDENTIFICADOR	TIPO DE REDE	SÍNTESE DA SIMULAÇÃO
Simulação 1	<i>Backpropagation</i>	Rede sendo utilizada como aproximador de funções;
Simulação 2	<i>Backpropagation</i>	Rede sendo utilizada como classificadora. Problema contém aprox. 40 classes;
Simulação 3	<i>Backpropagation</i>	Rede sendo utilizada como classificadora em problema simplificado de identificação de classes;
Simulação 4	BAM	Rede sendo utilizada como classificadora em problema simplificado de identificação de classes;
Simulação 5	Hopfield	Rede sendo utilizada como classificadora em problema simplificado de identificação de classes;
Simulação 6	GRNN	Rede sendo utilizada como classificadora em problema simplificado de identificação de classes;

4.10 – Perceptron multicamadas com algoritmo *backpropagation*

4.10.1 – Simulação 1

Devido às características do problema em questão, seria desejável que a rede pudesse identificar todos os tipos de curto e localizar a fase em que ocorreu o curto. Para adequar a rede a esse tipo de problema foi utilizada a estrutura de aproximador de funções que permite agregar a ela, uma maior quantidade de informações sobre o problema. No entanto, tal estrutura exige uma base de dados bastante representativa.

A arquitetura de uma rede neural com melhor desempenho para determinado problema é definida, normalmente, de forma empírica, porém, Chow, Mangum e Yee (1991) demonstram que: (a) estruturas de redes neurais de alta ordem apresentam maior precisão e menor tempo de treinamento; (b) a rede que utiliza maior número de nós na camada oculta

apresenta melhor desempenho. Assim, o espaço de entrada foi definido, originalmente, com 4 dimensões: $\{V, I_a, I_b, I_c\}$, conforme a TAB. 9 (capítulo 3), onde:

- V é a tensão média entre fases;
- I_a é a corrente medida na fase “A”;
- I_b é a corrente medida na fase “B”;
- I_c é a corrente medida na fase “C”.

Estas grandezas foram escolhidas a partir daquelas obtidas nas simulações descritas no capítulo 3. Foram feitas outras medidas como de potência e velocidade, porém as deficiências encontradas nos equipamentos utilizados tornaram tais medidas pouco confiáveis, por este motivo foram abandonadas.

Objetivando obter maior precisão e menor tempo de treinamento, o espaço de entrada foi expandido, acrescentando outras variáveis de entrada:

(a) o valor médio da corrente nas 3 fases, para cada tensão medida, durante os ensaios de curto;

(b) o produto entre o valor da tensão aplicada e a média dos valores das correntes sob curto - \bar{I} - obtendo outra grandeza: $V\bar{I}$;

(c) um outro parâmetro pôde ser obtido, interpolando os dados do ensaio a vazio, com frequência constante, em 60 Hz e tensões variáveis, para obter o valor da corrente a vazio – I_n , para a tensão de ensaio;

(d) o valor da corrente a vazio multiplicado pela tensão de ensaio V , gerando outro parâmetro: $V I_n$, correspondente à potência aparente no momento da simulação do curto.

Assim, o conjunto de padrões de entrada pôde ser expandido para 8 dimensões, conforme a TAB. 13: $\{V, I_n, I_a, I_b, I_c, \bar{I}, V I_n, V\bar{I}\}$.

O espaço de saída original $\{N_a, N_b, N_c\}$ permanece o mesmo, conforme a TAB. 9 apresentada no capítulo 3.

TABELA 13
Leiaute da base de dados expandida.

Ordem	Identificador	Significado
1 ^a	Vab	Valor eficaz da tensão de linha entre as fases.
2 ^a	In	Valor médio eficaz da Corrente nas 3 fases.
3 ^a	Ia	Valor eficaz da Corrente de linha na fase “A”.
4 ^a	Ib	Valor eficaz da Corrente de linha na fase “B”.
5 ^a	Ic	Valor eficaz da Corrente de linha na fase “C”.
6 ^a	\bar{I}	Média dos valores das correntes sob curto.
7 ^a	VIn	Potência obtida com a tensão e a corrente de ensaio, a vazio, sem curto-circuito.
8 ^a	$V\bar{I}$	Potência aparente, sob curto.
9 ^a	Na	N ^o . de espiras em curto na fase “A”.
10 ^a	Nb	N ^o . de espiras em curto na fase “B”.
11 ^a	Nc	N ^o . de espiras em curto na fase “C”.

4.10.1.1 – Estrutura da Rede *Backpropagation*

Foi utilizada uma rede com 8 nós na camada de entrada, 25 na camada escondida e 3 na camada de saída. Temos que levar em conta que 3 características definem o comportamento típico da *backpropagation*, a função de ativação dos nós da camada escondida, a função de ativação na camada de saída e o algoritmo utilizado para minimizar o erro médio quadrado. No presente caso temos:

1. **Função de ativação dos nós da camada escondida : SIGMÓIDE** pois ela é necessária para conseguirmos resolver problemas de características não-lineares, definida como “logsig” no MATLAB.
2. **Função de ativação dos nós na camada de saída – LINEAR** pois desejamos que a rede se comporte como um aproximador universal de funções.
3. **Algoritmo de otimização utilizado na minimização do erro médio quadrado –** Método determinístico de segunda ordem, baseado no método de Newton, denominado Levenberg-Marquardt.

Quanto ao particionamento da base de dados, de um total de 1126 padrões, 67% foi utilizado para treinamento e 33% dos dados para validação da rede.

4.10.1.2 – Implementação feita

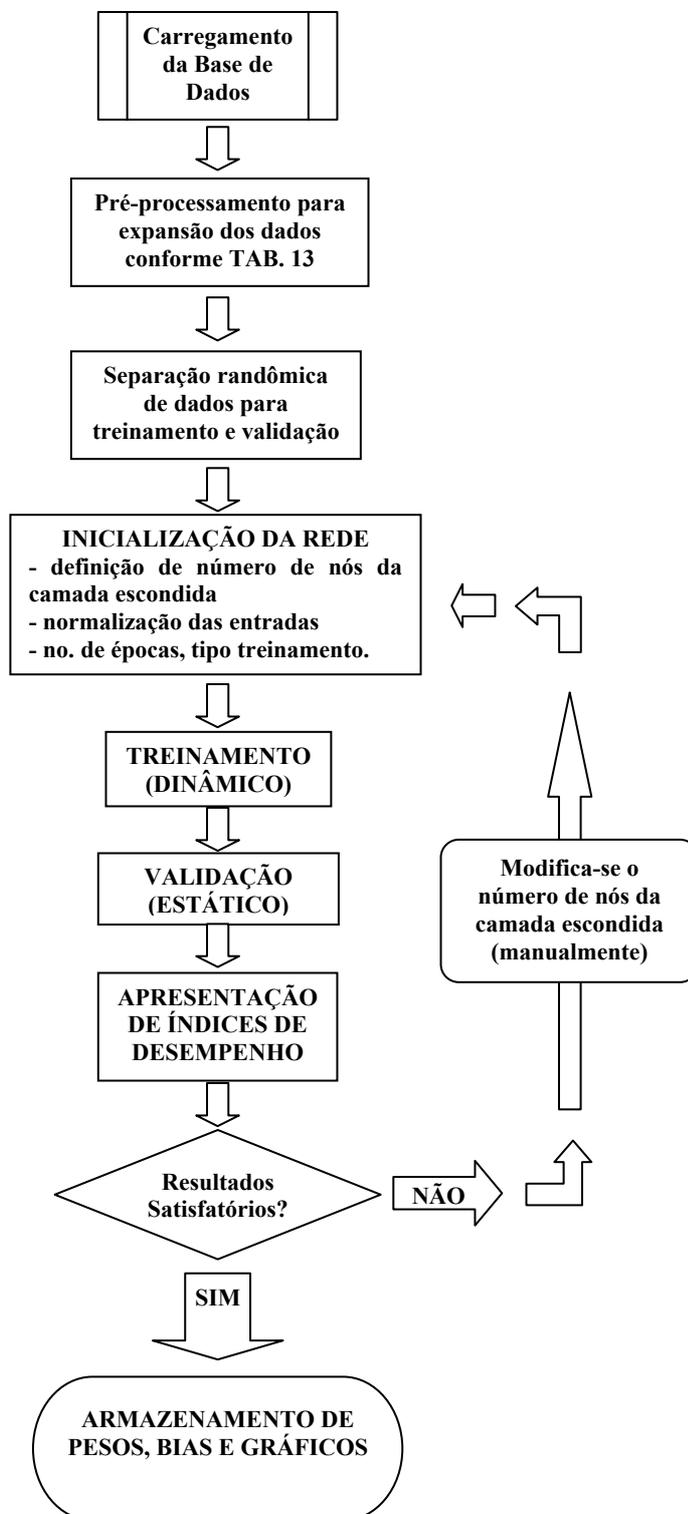
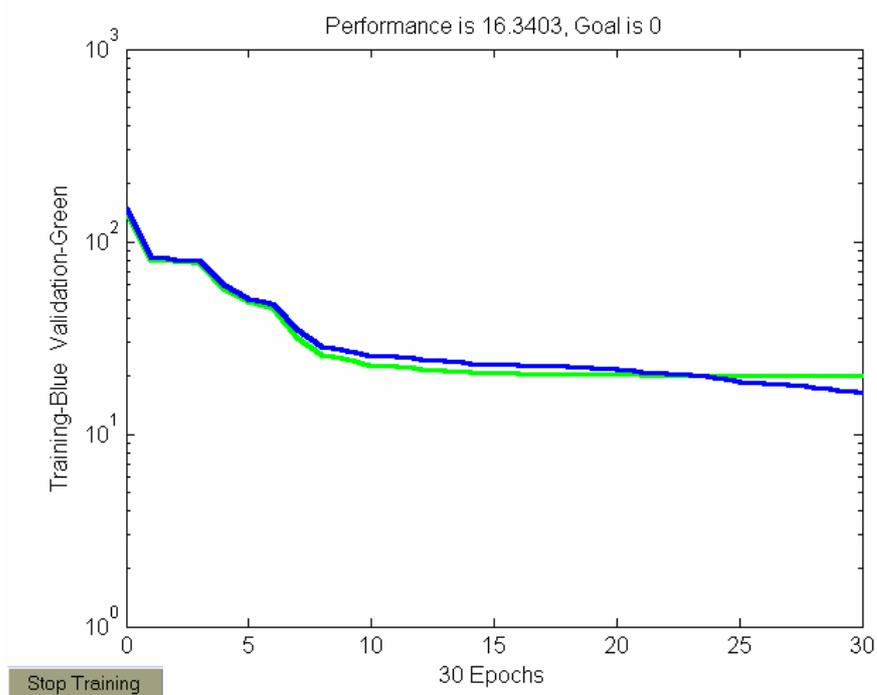


FIGURA 11 – Diagrama do processo de treinamento da rede na Simulação 1 (Rede *Backpropagation*).

4.10.1.3 – Resultados do Treinamento

O treinamento foi feito com parada antecipada. Na FIG. 12 pode-se visualizar a evolução do erro no treinamento. O erro utilizado foi o erro médio quadrado, conforme a equação (50) do APÊNDICE C e encontra-se em uma escala logarítmica de 1 a 1000 no gráfico. Os valores dos erros constituem-se em um somatório dos erros de cada padrão elevado ao quadrado, em cada iteração, originando ordenadas com valores elevados, diferentes de outras análises de desempenho em que o erro é dividido pelo número total de padrões. O erro de treinamento está em azul e o de validação em verde. A variável “*Performance*”, que se encontra no alto do gráfico, refere-se ao valor do erro médio quadrado no momento e o *goal* é o erro máximo tomado como objetivo.

FIGURA 12 – Evolução do treinamento da rede na Simulação 1 (Rede *Backpropagation*).

A evolução do treinamento da rede pode ser visualizada na FIG. 13, pela transcrição das mensagens que o MATLAB emitiu na tela da simulação, logo após a conclusão da simulação 1. São apresentados, respectivamente, a metodologia utilizada para minimização do erro, a quantidade de épocas (*Epoch*), o erro médio quadrado (*MSE*) e a norma do vetor gradiente (*Gradient*). O número de épocas máximo alcançado em cada ciclo de simulação e o número definido como objetivo, foram apresentados separados apenas por uma barra, o mesmo acontecendo com o erro médio quadrado e a norma do vetor gradiente. Nesse caso, o treinamento foi finalizado por parada antecipada:

```
TRAINLM, Epoch 0/50, MSE 149.436/0, Gradient 28819.5/1e-010  
TRAINLM, Epoch 5/50, MSE 50.7686/0, Gradient 9382.67/1e-010  
TRAINLM, Epoch 10/50, MSE 25.4622/0, Gradient 3404.81/1e-010  
TRAINLM, Epoch 15/50, MSE 23.0602/0, Gradient 1681.34/1e-010  
TRAINLM, Epoch 20/50, MSE 22.0128/0, Gradient 388.634/1e-010  
TRAINLM, Epoch 25/50, MSE 18.9033/0, Gradient 3778.7/1e-010  
TRAINLM, Epoch 30/50, MSE 16.3403/0, Gradient 636.818/1e-010  
TRAINLM, Validation stop.
```

FIGURA 13 – Trecho transcrito da tela de simulação 1, apresentando a metodologia de minimização de erros, o número de épocas, o erro médio quadrático, a norma do gradiente e a finalização da simulação.

4.10.1.4 – Validação

A validação foi feita utilizando outros dados diferentes daqueles utilizados para o ajuste de pesos, ou seja, esse teste tem por objetivo avaliar a capacidade de generalização da rede.

Nesse conjunto de testes existem 3 saídas que correspondem às espiras em curto nas fases A, B e C. A análise do erro não pôde ser feita em termos percentuais, pois estamos tratando de um problema de aproximação de uma função desconhecida e devido a isso não se

tem um padrão de comparação. São apresentados logo abaixo histogramas para facilitar a avaliação da qualidade do treinamento frente aos resultados da validação. Cada coluna do histograma representa o número de vezes em que o erro absoluto entre a saída desejada e a saída obtida se repetiu. Isso poder ser melhor compreendido, observando o exemplo mostrado na TAB. 14.

TABELA 14
Exemplo de como foram somadas as ocorrências de erro absoluto

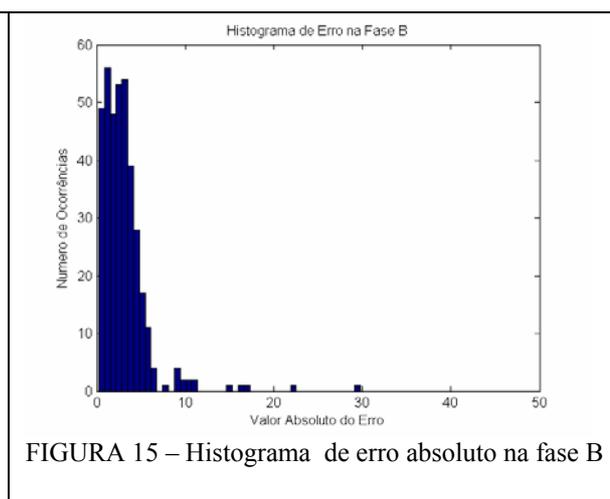
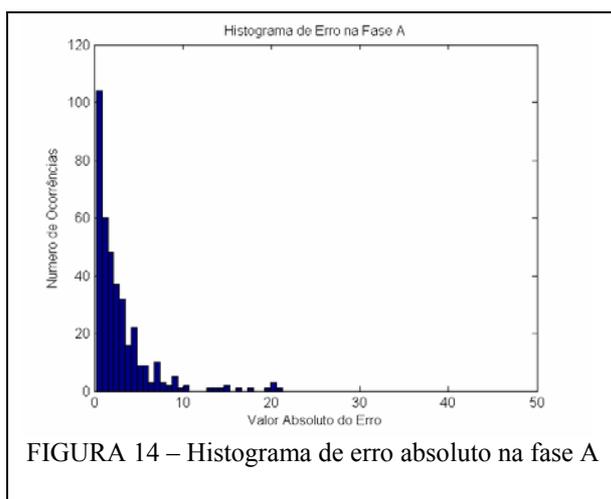
Identificação do Padrão	Saída desejada	Saída da rede	Erro absoluto	Ocorrência cumulativa do erro
X1	7	5	$= 7-5 = 2$	1
X5	50	52	$= 50-52 =2$	2
X6	18	18	$= 18-18 =0$	1
X7	32	32	$= 32-32 =0$	2
X4	0	0	$= 0-0 = 0$	3
X2	14	7	$= 14-7 = 7$	1
X3	10	5	$= 10 - 5 = 5$	1

TABELA 15
Distribuição do erro de validação para a simulação 1 com a Rede *Backpropagation*

Faixa do erro absoluto: saída desejada – saída obtida	Número de ocorrências do erro APENAS na Fase A, no intervalo da faixa	Número de ocorrências do erro APENAS na Fase B, no intervalo da faixa	Número de ocorrências do erro APENAS na Fase C, no intervalo da faixa	Número de ocorrências do erro global (nas 3 fases somadas para defeitos simultâneos) no intervalo da faixa
0.00 – 0.89	104	49	60	1
0.89 – 1.77	60	56	45	1
1.77 – 2.66	48	48	43	10
2.66 – 3.55	37	53	39	12
3.55 – 4.44	32	54	46	17
4.44 – 5.32	16	39	26	18
5.32 – 6.21	22	28	28	24
6.21 – 7.10	9	17	17	28
7.10 – 7.98	9	11	9	26

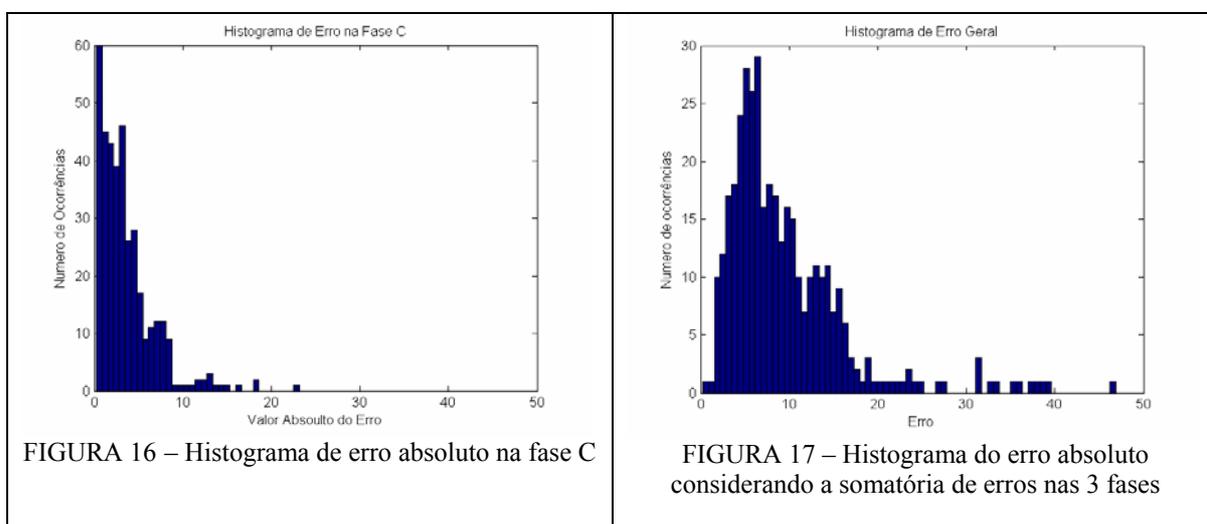
7.98 – 8.87	3	4	11	29
8.87 – 9.76	10	0	12	16
9.76 – 10.65	3	1	12	18
10.65 – 11.53	2	0	9	17
11.53 – 12.42	5	4	1	13
12.42 – 13.31	1	2	1	16
13.31 – 14.19	2	2	1	15
14.19 – 15.08	0	2	1	10
15.08 – 15.97	0	0	2	7
15.97 – 16.85	0	0	2	10
16.85 – 17.74	1	0	3	11
17.74 – 18.63	1	0	1	10
18.63 – 19.52	1	0	1	11
19.52 – 20.40	2	1	1	7
20.40 – 21.29	0	0	0	9
21.29 – 22.18	1	1	1	6
...

Se o número de amostras fosse infinito, o erro da rede seria determinado por uma distribuição gaussiana, com média μ e desvio padrão σ . Seria desejável que uma rede apresentando um treinamento satisfatório tivesse erro médio igual a zero e desvio padrão desprezível. Os gráficos das FIG. 14, 15, 16 e 17 são uma representação visual do conteúdo apresentado na TAB.15.



Na FIG. 14 é apresentado o histograma de erro com relação à primeira saída (NA). Podemos verificar que, para essa saída temos uma maior ocorrência do erro em valores próximos de zero e um desvio padrão considerável, observável através do perfil de distribuição do erro. Conclui-se que, na maioria das vezes, a rede acerta com relação à saída, porém o desvio padrão demonstra que há um espalhamento no erro, o que indica um desempenho não satisfatório.

A FIG. 15 apresenta o histograma de erro com relação à segunda saída (NB). Pode ser observada uma maior ocorrência do erro em valores próximos de duas unidades em relação ao valor de espiras esperado, ou seja, a rede não está conseguindo informar o valor correto com relação à essa saída, possivelmente devido à existência de padrões espúrios na massa de dados.



Na FIG. 16 é apresentado o histograma de erro com relação à terceira saída (NC). Essa saída apresentou o mesmo tipo de comportamento que a saída B. Na FIG. 17 é apresentado o histograma de erro global da rede, que possui uma maior ocorrência do erro em valores próximos de sete, considerando o erro total. Como demonstrado nos gráficos anteriores (FIG.

14, 15, 16 e 17), o treinamento da rede apresentou comportamento não satisfatório quando se consideram os resultados globais.

TABELA 16
Resultados da Simulação 1 com a Rede *Backpropagation*

Parâmetro de Desempenho	<i>Backpropagation</i> Simulação 1
• Tempo de treinamento (seg.)	129,67
• Tempo de validação (seg.)	0,18
• Função de ativação dos nós da camada escondida	Sigmóide
• Função de ativação dos nós da camada de saída	Linear
• Dimensão do vetor pesos da camada de entrada	25x8
• Dimensão do vetor pesos da camada de escondida	3x25
• Dimensão do vetor de <i>bias</i>	28x1
• Quantidade total de padrões na base de dados	1126
• Quantidade de padrões de treinamento	751
• Quantidade de padrões de validação	375
• Índice de Acerto no treinamento (em %)	Simulação suspensa (vide FIG. 12)
• Índice de Acerto na validação (em %)	Simulação suspensa (vide FIG. 17)

4.10.1.5 – Conclusões da Simulação 1 com a Rede *Backpropagation*

A FIG. 17 permite concluir que o problema não está bem modelado ou então a base de dados não apresenta uma quantidade de padrões significativos para tal tipo de estrutura.

A rede não conseguiu resolver o problema porque, para funcionar como aproximador de funções, a rede necessitaria de um número consideravelmente maior de padrões para ter comportamento satisfatório. O fato dos dados que compõem a saída não serem contínuos pode ser o responsável pela baixa qualidade obtida no treinamento. Por exemplo, supondo um caso

onde ocorrem 7 espiras em curto em uma determinada fase, não foram feitos ensaios com valores intermediários entre 0 e 7 espiras em curto, devido até mesmo, às dificuldades de se realizar tais ensaios com o protótipo físico utilizado.

4.10.2 – Simulação 2 (Rede *Backpropagation*)

Implementação semelhante à anterior com os mesmos tipos de entradas, variando-se a forma de representar as saídas. Já que as saídas contidas na base de dados apresentavam valores discretos, com grande espaçamento entre eles, foi desenvolvida uma nova estrutura baseado no seguinte critério de classificação: cada ensaio de curto-circuito diferente que compõe a base, passa a constituir uma classe.

Metodologia: foi realizado um pré-processamento para realizar o agrupamento das classes (clusterização), de acordo com o número de espiras em curto por fase, objetivando atribuir arbitrariamente a cada diferente tipo de ensaio realizado, uma classe, em função do número de espiras em curto e da localização das mesmas. Nessa clusterização foram identificadas 40 classes distintas, conforme apresentado na TAB. 8 (cap. 3).

Desse modo, a nova rede (Simulação 2) teve as saídas, anteriormente representadas por valores reais (NA, NB, NC) convertidas para **valores binários** que passaram a representar a classe ativada correspondente ao tipo e localização do curto-circuito.

O vetor de entradas foi mantido: $\{V_{ab}, I_n, I_a, I_b, I_c, \bar{I}, V_{In}, \bar{V}_I\}$.

Cada vetor de saída original $\{N_a, N_b, N_c\}$ foi convertido para binário e codificado, utilizando-se 6 bits.

TABELA 17
Leiaute da base de dados expandida.

Ordem	Identificador	Significado
1 ^a	Vab	Valor eficaz da tensão de linha entre as fases.
2 ^a	In	Valor médio eficaz da Corrente nas 3 fases.
3 ^a	Ia	Valor eficaz da Corrente de linha na fase “A”.
4 ^a	Ib	Valor eficaz da Corrente de linha na fase “B”.
5 ^a	Ic	Valor eficaz da Corrente de linha na fase “C”.
6 ^a	\bar{I}	Média dos valores das correntes sob curto.
7 ^a	VIn	Potência obtida com a tensão e a corrente de ensaio, a vazio, sem curto-circuito.
8 ^a	$V\bar{I}$	Potência aparente, sob curto.
9 ^a	6 bits	Codificação das 40 classes diferentes de curto-circuito simuladas, utilizando 6 bits.

4.10.2.1 – Estrutura da Rede

Foi utilizada uma rede com 8 nós na camada de entrada, 25 na camada escondida e 6 na camada de saída. As características estruturais da rede são as seguintes:

1. **Função de ativação dos dos nós da camada escondida** : **SIGMÓIDE** pois ela é necessária para conseguirmos resolver problemas de características não-lineares, definida como “logsig” no MATLAB.
2. **Função de ativação dos nós na camada de saída** – **SIGMÓIDE** pois desejamos que a rede se comporte como um classificador, ou seja, a saída contém a informação codificada de qual classe foi ativada.

3. Algoritmo de otimização utilizado na minimização do erro médio quadrado –

Método determinístico de segunda ordem, baseado no método de Newton, denominado Levenberg-Marquardt.

Como a saída da rede proposta encontra-se em formato binário, a saída da sigmóide é contínua e está presente no intervalo [0,1]. Sabendo-se desse fato, arbitrou-se um limiar acima do qual, os valores serão convertidos para “1” e abaixo do qual, serão convertidos para “0”. O limiar escolhido foi 0.5 e a implementação dessa estratégia foi possível através do uso de uma função de arredondamento do MATLAB -“round”- na saída avaliada da rede.

Características gerais:

- Quantidade total de padrões na base de dados: 1126.
 - Quantidade de padrões utilizados no treinamento: 751.
 - Quantidade de padrões utilizados para fazer a validação: 375.
-

4.10.2.2 – Implementação feita

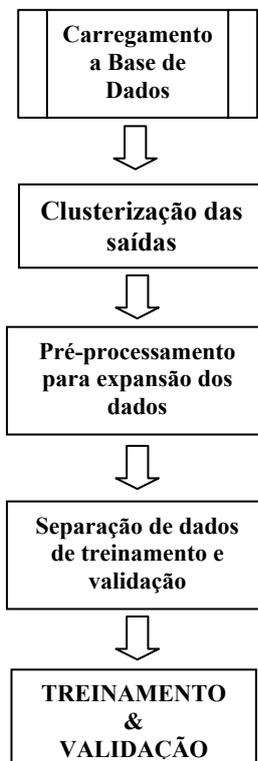


FIGURA 18 – Diagrama do processo de treinamento da rede na Simulação 2.

4.10.2.3 – Resultados do Treinamento

Foi utilizado o recurso conhecido como parada antecipada, assim como na **Simulação 1**. Pode ser avaliada a evolução dos erros de treinamento e validação através da FIG. 19. Os seguintes percentuais foram fornecidos pelo programa de simulação:

- Índice de acerto no treinamento: 62,31%.
- Índice de acerto na validação: 57,06%.

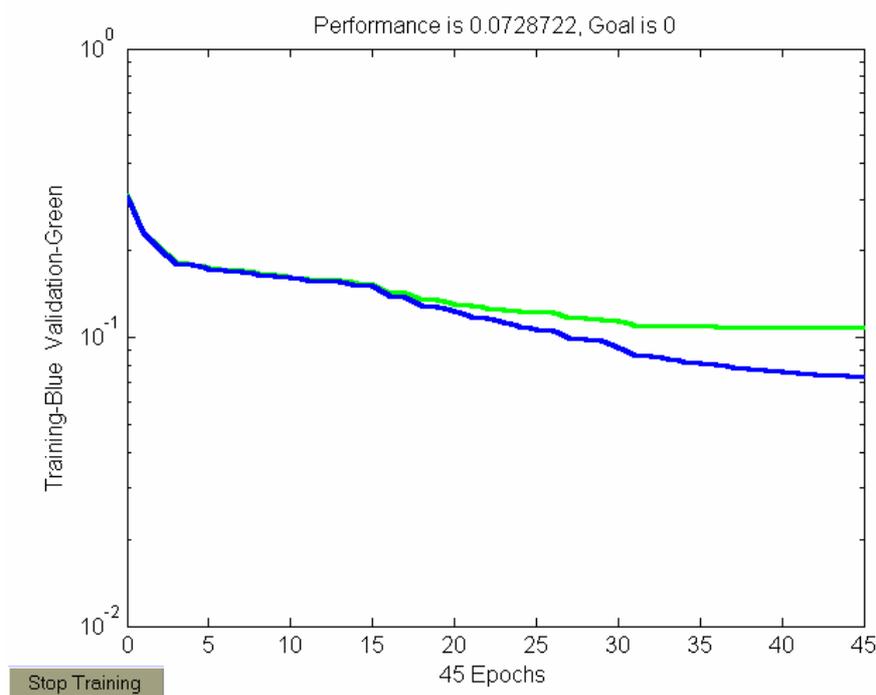


FIGURA 19 – Evolução do treinamento da rede na simulação 2.

A evolução do treinamento desta rede pode ser visualizada pelo trecho transcrito da tela de simulação, na FIG. 20, apresentando, por ciclo de simulação, a metodologia para minimização do erro, a quantidade de épocas, o erro médio quadrado e a norma do gradiente, estando indicados os valores alcançados, bem como aqueles definidos como objetivos. Treinamento finalizado por parada antecipada.

```

TRAINLM, Epoch 0/50, MSE 0.306509/0, Gradient 165.905/1e-010
TRAINLM, Epoch 5/50, MSE 0.173359/0, Gradient 13.942/1e-010
TRAINLM, Epoch 10/50, MSE 0.161311/0, Gradient 28.0805/1e-010
TRAINLM, Epoch 15/50, MSE 0.150826/0, Gradient 117.401/1e-010
TRAINLM, Epoch 20/50, MSE 0.123185/0, Gradient 40.0972/1e-010
TRAINLM, Epoch 25/50, MSE 0.107464/0, Gradient 12.1139/1e-010
TRAINLM, Epoch 30/50, MSE 0.0927645/0, Gradient 70.5908/1e-010
TRAINLM, Epoch 35/50, MSE 0.0814181/0, Gradient 61.8915/1e-010
TRAINLM, Epoch 40/50, MSE 0.0762569/0, Gradient 5.52053/1e-010
TRAINLM, Epoch 45/50, MSE 0.0728722/0, Gradient 26.9089/1e-010
TRAINLM, Validation stop.

```

FIGURA 20 – Trecho transcrito da tela de simulação 2, apresentando a metodologia de minimização de erros, o número de épocas, o erro médio quadrático, a norma do gradiente e a finalização da simulação.

4.10.2.4 – Validação

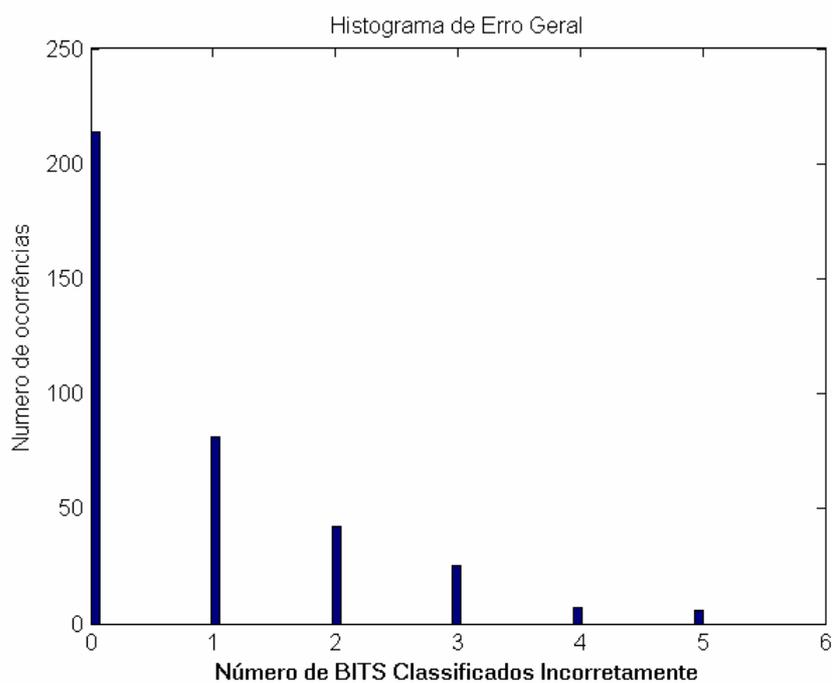


FIGURA 21 – Histograma de erro global de desempenho - Simulação 2.

A FIG. 21 representa um histograma do erro global na classificação gerada sobre os dados de validação. Cada número no eixo das abcissas representa a quantidade de bits que foram identificados de forma errada, simultaneamente. O eixo das ordenadas apresenta a quantidade de vezes em que o número de bits da abscissa falhou, cumulativamente, durante a validação, comparando-se a saída desejada com a real.

TABELA 18
Resultados da simulação 2 com a Rede *Backpropagation*

Parâmetro de Desempenho	<i>Backpropagation</i> Simulação 2
• Tempo de treinamento (seg.)	730,50
• Tempo de validação (seg.)	0,12
• Função de ativação dos nós da camada escondida	sigmóide
• Função de ativação dos nós da camada de saída	sigmóide
• Dimensão do vetor pesos da camada de entrada	25x8
• Dimensão do vetor pesos da camada de escondida	6x25
• Dimensão do vetor de <i>bias</i>	31x1
• Quantidade total de padrões na base de dados	1126
• Quantidade de padrões de treinamento	751
• Quantidade de padrões de validação	375
• Índice de Acerto no treinamento (em %)	62,31
• Índice de Acerto na validação (em %)	57,06

4.10.2.5 – Conclusões da Simulação 2 com a Rede *Backpropagation*

O significativo aumento no tempo de treinamento e o elevado índice de erro de validação indicam que a rede não conseguiu resolver o problema, possivelmente devido à sua complexidade. Um dos motivos supostos para explicar esse fato foi o de que a massa de dados não era realmente significativa para treinar essa quantidade tão grande de classes. Algumas classes apresentaram número de padrões inferior a 10. A inserção das colunas relativas às potências não proporcionou uma melhoria no treinamento da rede, possivelmente por serem linearmente dependentes das outras colunas e não apresentarem informações novas na base de dados.

4.10.3 – Simulação 3 com a Rede *Backpropagation*

Essa última estratégia teve por objetivo simplificar, ao máximo, o problema para que fosse possível para a rede, a partir de uma massa de dados pequena, conseguir chegar a resultados satisfatórios.

Nesta modelagem:

- (a) não se pretende identificar o número de espiras em curto-circuito por fase, mas tão somente identificar se ocorreu ou não, curto em cada uma das 3 fases;
- (b) deseja-se classificar os curtos segundo a gravidade dos mesmos.

A base de dados recebeu um pré-processamento no qual foi inserido, temporariamente, um campo com a corrente a vazio, sob condições normais de funcionamento, para a tensão de ensaio, considerando-se a interpolação feita conforme visto no capítulo 4 (FIG. 8). Logo depois, foi criado um índice para indicar a gravidade dos curtos, tomando-se, como índice, a divisão da corrente média eficaz das 3 fases pela corrente a vazio, sem curto, em cada tensão de ensaio:

$$\frac{\bar{I}}{I_n} = a \quad (11)$$

As seguintes faixas de gravidade de curto-circuito foram tomadas de forma empírica, tendo em vista a experiência pessoal:

- $a < 1.3$ → curto incipiente ou situação sem curto;
- $1.3 \leq a < 1.6$ → curto médio
- $a \geq 1.6$ → curto grave

A seguir, cada ensaio foi codificado e a classificação inserida na base de dados, utilizando-se código binário, de acordo com:

- “0 0” para representar situação sem curto ou curto incipiente;
- “0 1” para representar curto médio;
- “1 1” para curto grave.

As colunas que representavam valores de potência do sistema também foram excluídas para simplificar o treinamento e retirar dados linearmente dependentes. Sendo assim, as saídas passaram a ser representadas por 5 bits, de modo que os 3 primeiros informam em qual fase ocorreu o curto e os 2 bits restantes identificam a sua gravidade. A base de dados apresenta então o leiaute mostrado na TAB. 19.

O vetor de entradas: $\{V_{ab}, I_n, I_a, I_b, I_c\}$.

O vetor de saída: tipos de curtos-circuitos convertidos para binário e codificados, utilizando-se 5 bits.

TABELA 19
Leiaute da base de dados expandida.

Ordem	Identificador	Significado
1 ^a	Vab	Valor eficaz da tensão de linha entre as fases.
2 ^a	In	Valor médio eficaz da Corrente nas 3 fases.
3 ^a	Ia	Valor eficaz da Corrente de linha na fase “A”.
4 ^a	Ib	Valor eficaz da Corrente de linha na fase “B”.
5 ^a	Ic	Valor eficaz da Corrente de linha na fase “C”.
6 ^a	5 bits	Codificação simplificada dos tipos de curtos possíveis

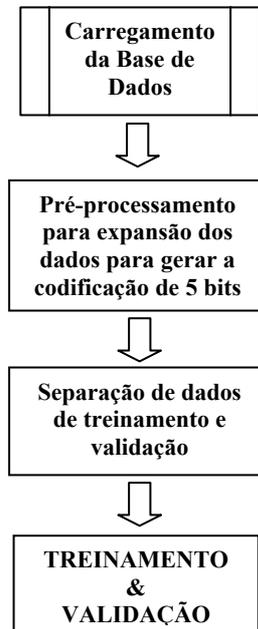
A saída do problema proposto se encontra em formato binário, entretanto, a saída da sigmóide é contínua e está presente no intervalo $[0,1]$. Considerando-se esse fato, foi arbitrado um limiar tal que, qualquer valor que se encontre acima dele, será convertido para “1” e os valores abaixo, serão convertidos para “0”. O limiar escolhido foi 0.5 e a implementação dessa estratégia foi possível utilizando a função de arredondamento “round” na saída avaliada da rede.

4.10.3.1 – Estrutura (Simulação 3 com a Rede *Backpropagation*):

Foi utilizada uma rede com 5 nós na camada de entrada, 15 na camada escondida e 5 na camada de saída. As características estruturais da rede são as seguintes:

1. **Função de ativação dos dos nós da camada escondida : SIGMÓIDE** pois ela é necessária para conseguirmos resolver problemas de características não-lineares, definida como “logsig” no MATLAB.
 2. **Função de ativação dos nós na camada de saída – SIGMÓIDE** pois desejamos que a rede se comporte como um classificador, ou seja, a saída contém a informação codificada de qual classe foi ativada.
 3. **Algoritmo de otimização utilizado na minimização do erro médio quadrado –** Método determinístico de segunda ordem, baseado no método de Newton, denominado Levenberg-Marquardt.
-

4.10.3.2 – Implementação feita

FIGURA 22 – Diagrama do processo de treinamento da rede na Simulação 3 com a Rede *Backpropagation*.

4.10.3.3 – Resultados do Treinamento

Foi utilizado treinamento incluindo o recurso da parada antecipada. Pode ser avaliada a evolução dos erros de treinamento e validação na FIG. 23.

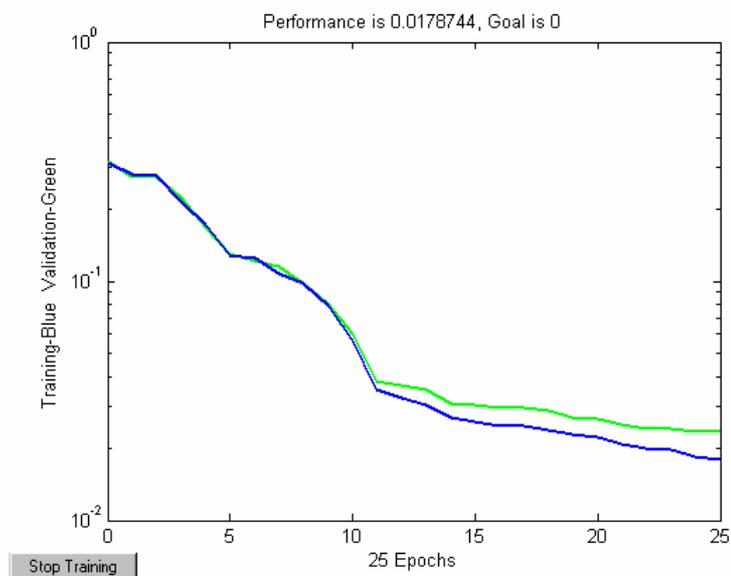


FIGURA 23 – Evolução do treinamento da rede na simulação 3

Na FIG. 24 pode-se visualizar a evolução do treinamento da rede bem como a metodologia para minimização do erro utilizada, a evolução em épocas, o erro médio quadrado e a norma do gradiente:

```
TRAINLM, Epoch 0/25, MSE 0.355415/0, Gradient 210.989/1e-010  
TRAINLM, Epoch 5/25, MSE 0.234496/0, Gradient 160.375/1e-010  
TRAINLM, Epoch 10/25, MSE 0.180614/0, Gradient 9.92358/1e-010  
TRAINLM, Epoch 15/25, MSE 0.122191/0, Gradient 138.406/1e-010  
TRAINLM, Epoch 20/25, MSE 0.028324/0, Gradient 41.0372/1e-010  
TRAINLM, Epoch 25/25, MSE 0.0183952/0, Gradient 8.9731/1e-010  
TRAINLM, Maximum epoch reached, performance goal was not met.
```

FIGURA 24 – Trecho transcrito da tela de simulação 2, apresentando a metodologia de minimização de erros, o número de épocas, o erro médio quadrático, a norma do gradiente e a finalização da simulação.

4.10.3.4 – Validação

TABELA 20
Desempenho Estatístico da Simulação 3 da Rede *Backpropagation*

Codificação	ACERTOS	ERROS
Bit 1 (curto individual na fase A)	358 (95,47 %)	17
Bit 2 (curto individual na fase B)	368 (98,13 %)	7
Bit 3 (curto individual na fase C)	343 (91,47 %)	32
Bit 4 (gravidade do curto - individual)	372 (99,20 %)	3
Bit 5 (gravidade do curto - individual)	374 (99,73 %)	1
Considerando todos os Bits simultaneamente	328 (87,47 %)	47

Na TAB.20 podemos observar que o elevado número de acertos para todos os bits tomados individualmente (Bits 1 a 5) demonstram que esta rede (Simulação 3 da Rede *Backpropagation*) é capaz de identificar se existiu ou não um curto nas Fases A, B e C, bem como informar se o curto é grave ou não. O percentual de acerto ainda continua aceitável se considerarmos o acerto de todos os bits, simultaneamente (87,47 %).

A FIG. 25 complementa a análise, apresentando a ocorrência de erros quando se considera até 2 bits com erros simultâneos.

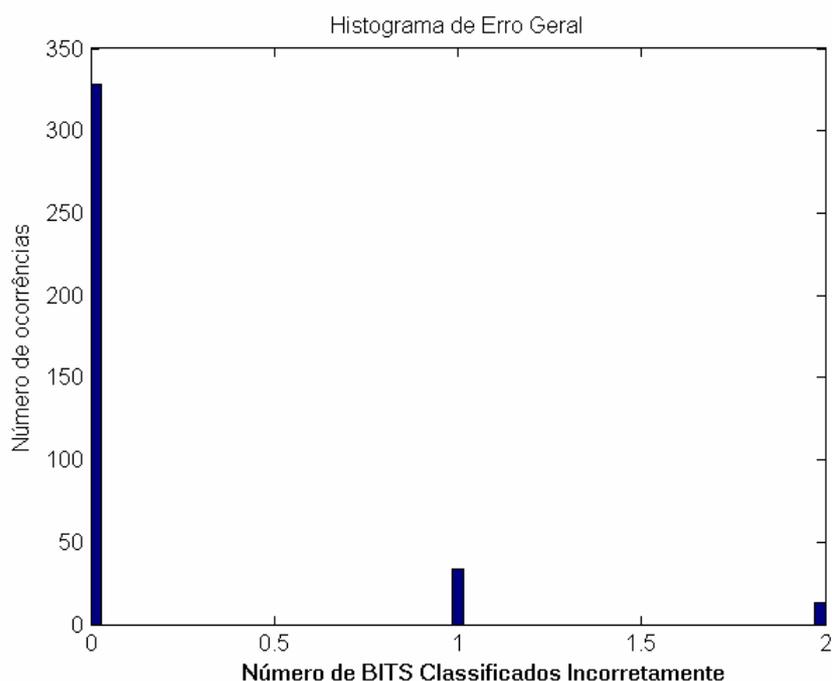


FIGURA 25 – Histograma de erro global

A FIG. 25 representa um histograma de erro global na classificação, onde o número fracionário não tem significado, tendo sido usado pelo MATLAB para compor o histograma e apresentá-lo em tamanho compatível. As abscissas indicam a quantidade de bits errados, independentemente de quais foram. As ordenadas indicam o número de vezes em que a rede cometeu tais erros.

TABELA 21
Resultados da simulação 3 com a Rede Backpropagation

Parâmetro de Desempenho	Backpropagation Simulação 3
• Tempo de treinamento (seg.)	65,41
• Tempo de validação (seg.)	0,27
• Função de ativação dos nós da camada escondida	sigmóide
• Função de ativação dos nós da camada de saída	sigmóide
• Dimensão do vetor pesos da camada de entrada	15x5
• Dimensão do vetor pesos da camada de escondida	5x15

• Dimensão do vetor de bias	20x1
• Quantidade total de padrões na base de dados	1126
• Quantidade de padrões de treinamento	751
• Quantidade de padrões de validação	375
• Índice de Acerto no treinamento (em %)	90,54
• Índice de Acerto na validação (em %)	87,46

4.10.3.5 – Conclusões da Simulação 3 da Rede *Backpropagation*

Pode-se visualizar, tanto graficamente quanto na TAB. 21, que o desempenho da rede foi satisfatório, superando todas as demais simulações e até mesmo as diversas topologias pesquisadas (resultados serão apresentados nas seções seguintes). É curioso notar que até mesmo o tamanho da estrutura da rede pôde ser reduzido, o que também pode ser encarado como ponto positivo para essa simulação, pois foi reduzida a complexidade da rede.

A rede conseguiu resolver o problema, mas é fato que, para casos em que exista treinamento *on-line*, deve-se ponderar a respeito de sua utilização, como no caso de aplicações em tempo real. Neste tipo de aplicação, as ações de controle devem ser tomadas em milissegundos, o que inviabiliza o treinamento, pois ele gasta uma quantidade significativa de memória e seu tempo é relativamente alto. Habitualmente, para resolver este problema, é feito um treinamento em paralelo e em seguida é feita a atualização dos pesos que estão sendo utilizados em alguma aplicação, desse modo não se interfere tanto na dinâmica do processo.

Nos casos em que o treinamento *on-line* não é necessário, comprovamos que tal topologia é adequada. A avaliação da rede é feita de uma forma muito rápida, pois se trata de simples multiplicações matriciais, esse é um ponto positivo do MATLAB, utilizado como

ferramenta para criar e simular as redes. No caso da rede *backpropagation*, considere-se também que seu treinamento é dinâmico e lento, mas sua avaliação é estática e rápida.

4.11 – Memória Associativa Bidirecional (BAM)

4.11.1 – Simulação 4 (BAM)

Durante o desenvolvimento foram realizados 3 experimentos com a rede BAM, de forma equivalente aos procedimentos efetuados com a *backpropagation* (seções 4.10.1, 4.10.2 e 4.10.3). Será apresentado apenas o ensaio que obteve melhores resultados, que corresponde à metodologia descrita na seção 4.10.3 e com formato da base de dados baseado na TAB. 19. O algoritmo utilizado no treinamento das diferentes configurações da rede BAM foi o mesmo utilizado em todas as simulações, tendo sido alterados somente o tipo de dados carregados para treinamento e validação. A BAM recebe tais dados já devidamente processados, expandidos e normalizados.

O pré-processamento, feito em etapa anterior ao treinamento, utiliza o método de Gram-Schmidt para geração de bases ortogonais, objetivando retirar dos dados de treinamento, os vetores linearmente dependentes. O treinamento da BAM fica favorecido se for utilizado algum método para ortonormalização dos dados. O método QR de decomposição matricial, recurso do MATLAB, realiza a fatoração QR, assim:

$$\mathbf{X} = \mathbf{Q}\mathbf{R} \tag{12}$$

onde:

$\mathbf{Q} \rightarrow$ matriz ortonormal resultante da decomposição. Sendo \mathbf{Q} ortonormal, a transposta de $\mathbf{Q}(\mathbf{Q}^T)$ corresponde à sua própria inversa \mathbf{Q}^{-1} ;

$\mathbf{R} \rightarrow$ matriz triangular superior resultante do escalonamento de \mathbf{X} , ou seja, todos os elementos abaixo da diagonal de \mathbf{R} são nulos.

A equação de treinamento da BAM pode ser vista a seguir:

$$\mathbf{M} = \mathbf{Y} \cdot \mathbf{Q} \quad (13)$$

$$\mathbf{Y} = \mathbf{M} \cdot \mathbf{Q}^{-1} \quad (14)$$

Da equação (12) temos:

$$\mathbf{X} \cdot \mathbf{R}^{-1} = \mathbf{Q} \cdot \mathbf{R} \cdot \mathbf{R}^{-1} \quad (15)$$

$$\mathbf{X} \cdot \mathbf{R}^{-1} = \mathbf{Q} \quad (16)$$

$$\mathbf{Q}^{-1} = \mathbf{Q}^T = (\mathbf{X} \cdot \mathbf{R}^{-1})^T \quad (17)$$

Substituindo-se equação (17) em equação (14), temos a equação de recuperação do valor presente na memória:

$$\mathbf{Y} = \mathbf{M} \cdot (\mathbf{X} \cdot \mathbf{R}^{-1})^T \quad (18)$$

4.11.1.1 – Estrutura da Rede (BAM)

Foram utilizados 250 bits para codificar os vetores de entrada, ou seja, 50 bits para cada entrada. Tendo em vista esse fato, a estrutura definida para a rede consiste em:

- 1 matriz quadrada \mathbf{Q} formada de vetores ortonormais (obtidos da base binarizada) entre si, tal matriz é de dimensão $\mathbf{L} \times \mathbf{L}$ onde \mathbf{L} é o número de padrões em \mathbf{X} ;
 - 1 matriz triangular superior \mathbf{R} que possui a mesma dimensão de \mathbf{X} ;
-

- 1 matriz \mathbf{M} que tem dimensão igual à transposta da matriz de dados linearmente independentes gerada no pré-processamento;

A quantidade total de padrões que compõem a base de dados é de 1126 padrões, no entanto, foi necessário ortonormalizar tal base, o que reduziu as quantidades de padrões para treinamento e validação da BAM.

- Quantidade de padrões de treinamento: 188.
- Quantidade de padrões de validação: 375.

A quantidade de padrões reduzida na fase de treinamento para a BAM, em relação ao número de padrões utilizados nas outras simulações, é devida ao fato da BAM ter apresentado inconsistência numérica para matrizes de entrada com dimensões muito elevadas. As alternativas para tentar armazenar mais dados seria a de aumentar o número de bits, porém essa estratégia implica na formação de matrizes mal condicionadas e não inversíveis, quando tais matrizes possuísem muitas colunas linearmente dependentes.

4.11.1.2 – Implementação feita

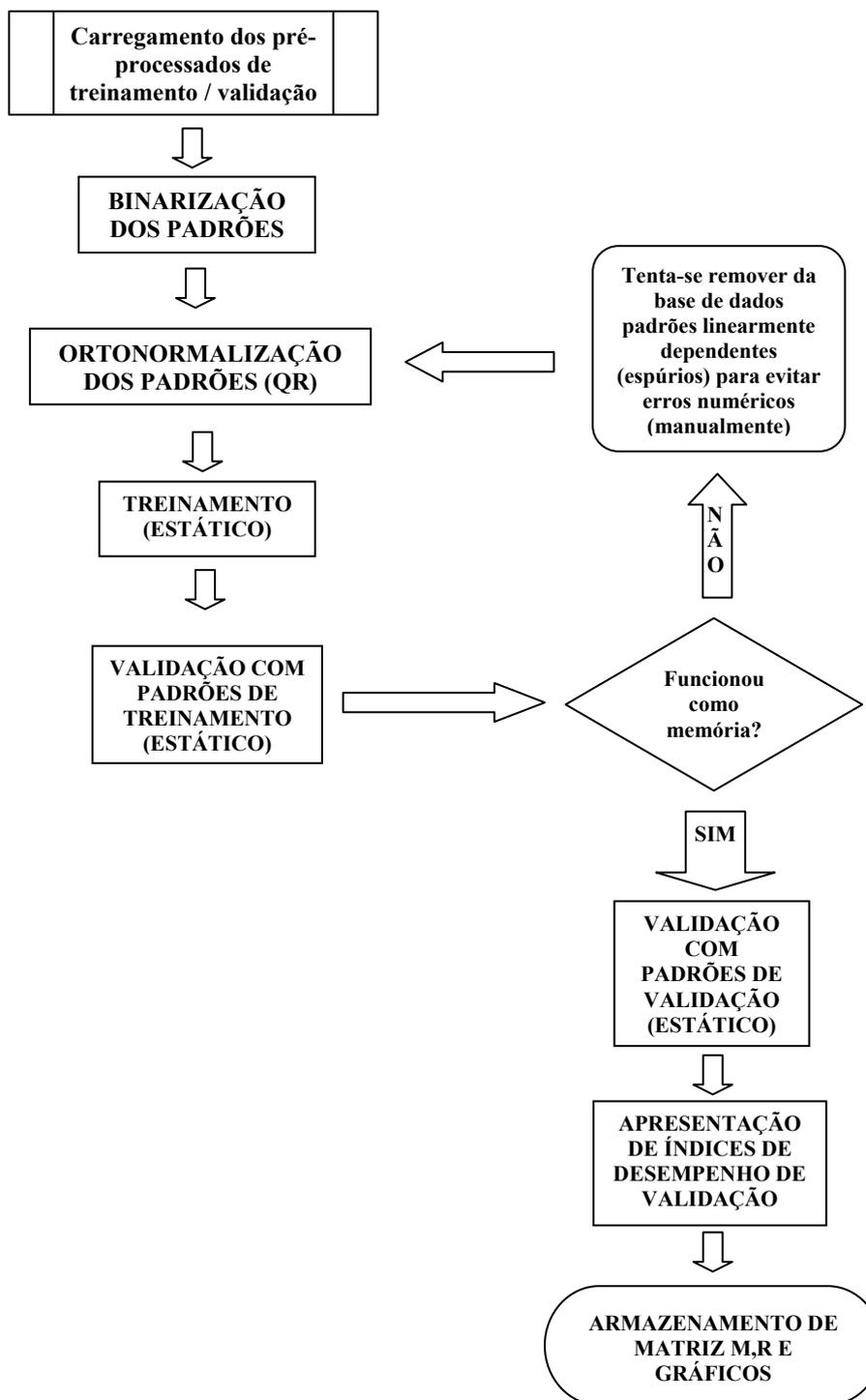


FIGURA 26 – Diagrama do processo de treinamento da BAM

4.11.1.3 – Resultados do Treinamento da BAM

Seguem-se os resultados de treinamento da BAM. Sua avaliação como memória associativa deve levar em conta a capacidade de generalização, testada com o procedimento de validação.

Os seguintes percentuais foram fornecidos pelo programa de simulação:

- Índice de acerto no treinamento: 100%.
- Índice de acerto na validação: 13,06%.

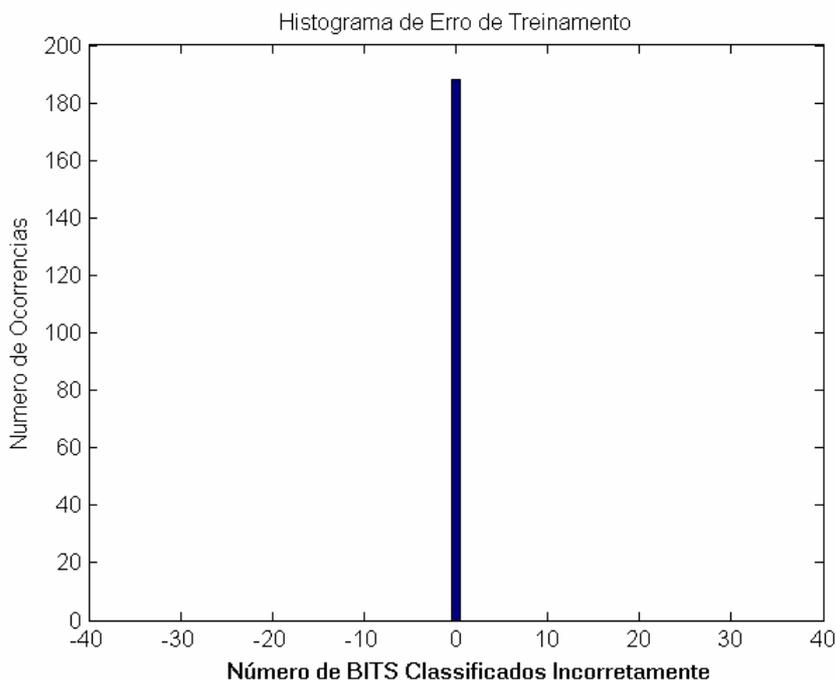


FIGURA 27 – Histograma de Erro no Treinamento da BAM.

Como pode ser observado na FIG. 27, o desempenho da rede frente aos padrões de treinamento, foi excelente, com 100% de acerto. Pode ser verificado que, em termos de

memória, a BAM apresentou um desempenho muito satisfatório, conseguindo armazenar os padrões a ela submetidos, durante a fase de treinamento.

4.11.1.4 – Validação da BAM

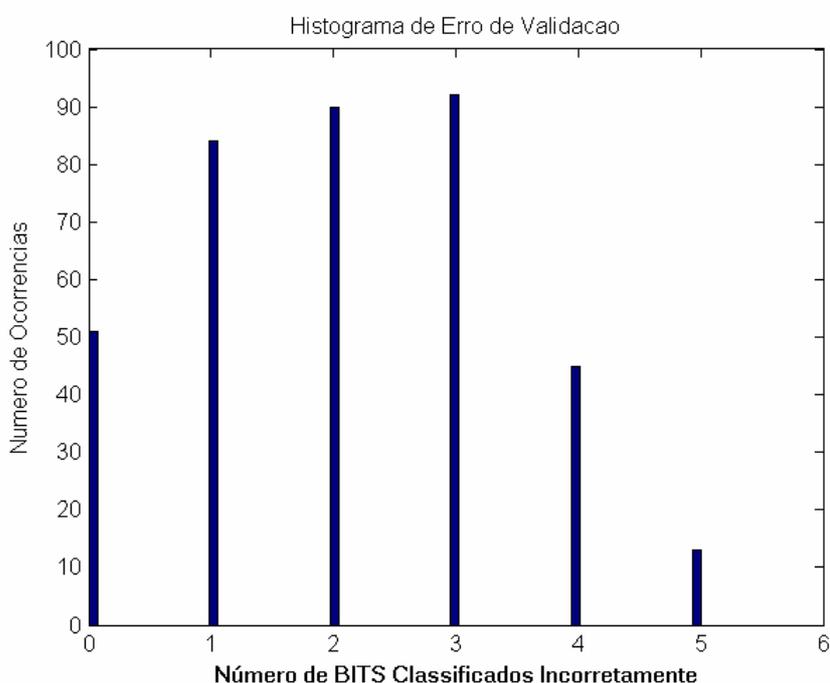


FIGURA 28 – Histograma de Erro na Validação

A capacidade de generalização – APÊNDICE C– ficou comprometida, já que, durante a validação (FIG. 28) a BAM utilizada não conseguiu apresentar resultados satisfatórios: apenas cerca de 50 padrões foram classificados corretamente. O que essa rede permite realizar é somente convergência para um padrão mais próximo dentro do universo de dados para o qual foi treinada.

TABELA 22
Resultados da Rede BAM

Parâmetro de Desempenho	BAM
• Tempo de treinamento (seg.)	1,1500
• Tempo de validação (seg)	0,9300
• Matriz M	5x188
• Matriz Q	188x188
• Matriz R	188x250
• Quantidade total de padrões na base de dados	1126
• Quantidade de padrões de treinamento	188
• Quantidade de padrões de validação	375
• Índice de Acerto no treinamento (em %)	100,0
• Índice de Acerto na validação (em %)	13,60

4.11.1.5 – Conclusões da Simulação 4

Pode-se ressaltar que a BAM possui um treinamento e uma avaliação muito rápidos, mas isso pode, em algumas situações, mascarar o intenso processamento da base de dados necessário para que se obtenha resultados satisfatórios. Os dados reais obtidos a partir dos experimentos feitos no protótipo (Capítulo 3) precisaram ser passados para formato binário, além de receberem tratamento para garantir a independência linear dos padrões de entrada, para que o método QR funcionasse corretamente. No entanto, os problemas de processamento de imagens, pelo fato de já possuírem uma representação dos dados em formato binário, utilizam preferencialmente esse tipo de topologia.

Uma estratégia para melhorar o desempenho da BAM pode ser o de aumentar o número de bits para codificação das entradas, isso também gera um aumento de tempo no pré-processamento e, dependendo do tamanho da base de dados empregada no treinamento, pode desencorajar seu uso, visto que outras topologias trabalham com dados reais, sem a necessidade de um pré-processamento tão intenso.

A incapacidade de generalização, ou seja, de extrapolar ou mesmo interpolar alguma tendência na base de dados de treinamento, faz da BAM uma alternativa não muito interessante para o presente problema. Tal afirmação pode ser comprovada pelo baixo índice de acerto na validação, apresentado na TAB. 22.

4.12 – Rede Hopfield

4.12.1 – Simulação 5

Durante o desenvolvimento foram realizados 3 experimentos com a rede Hopfield, equivalentes às simulações feitas utilizando a rede *backpropagation* (seções 4.10.1, 4.10.2 e 4.10.3). No entanto, será apresentado apenas o ensaio que obteve melhores resultados e que corresponde à metodologia descrita na seção 4.10.3, com formato da base de dados baseado na TAB. 19.

4.12.1.1 – Estrutura da Rede

A rede Hopfield possui um treinamento estático, ou seja, os seus pesos são definidos por regras sem utilização de algum aprendizado que envolva alguma dinâmica (iterações). A escolha dos padrões de treinamento também difere das demais redes. Visto que quanto mais

dados espúrios existirem no treinamento, pior será o desempenho da rede, diferentemente da *Backpropagation* e da GRNN pois tais redes se baseiam no fato de que, quanto maior a quantidade de dados, mais favorecidas tornam-se as redes para generalizar. O treinamento da Hopfield consiste em separar da base de dados, atratores singulares que representem tipos de padrões para que possa haver convergência quando forem submetidos padrões semelhantes. Nesta convergência, padrões semelhantes sofreriam atração por um mesmo atrator.

Como a Hopfield possui uma estrutura onde a entrada é igual à saída, foi feito um artifício de modo a conseguir adaptar o problema à topologia em questão: um padrão de entrada possui, ao mesmo tempo, entradas e saídas concatenadas. A idéia é de que, conseguindo-se um atrator que representasse corretamente um determinado tipo de falha, o vetor com as entradas fosse atraído, convergindo para o vetor com o qual tivesse maior afinidade. O atrator que fosse capaz de atrair o vetor de entrada forneceria as saídas a serem comparadas com os valores esperados.

Na fase de pré-processamento, os dados foram convertidos em valores bipolares e, em seguida, concatenou-se a entrada de cada padrão com sua saída esperada, compondo o vetor de entrada, conforme a FIG.29-A.

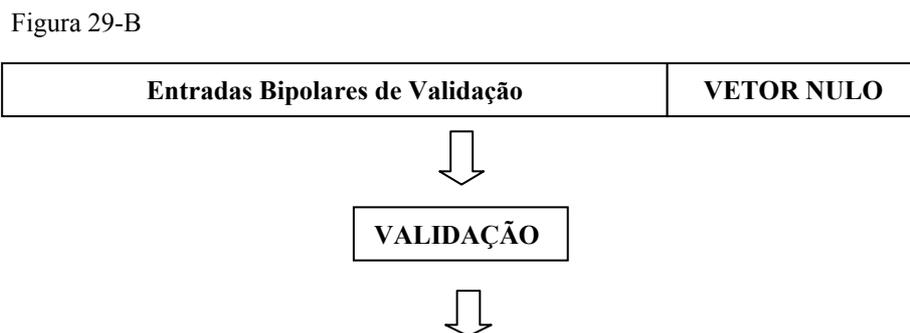
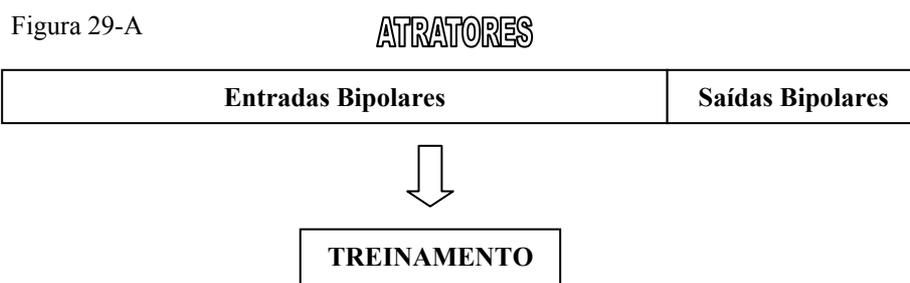


FIGURA 29 – Processo de seleção dos atratores e avaliação dos padrões para a rede Hopfield.

Na fase de validação (FIG.29-B), as saídas bipolares foram substituídas por vetores nulos, objetivando que os padrões de validação apresentados fossem atraídos pelos atratores que apresentassem maior similaridade.

A FIG.29-C sintetiza o que ocorre quando os padrões de validação são atraídos pelos atratores: as saídas bipolares do atrator são encontradas e assimiladas, encontrando assim, a resposta da rede aos padrões incompletos apresentados para validá-la.

4.12.1.2 – Implementação feita

Foi utilizado o algoritmo padrão do MATLAB para treinamento da rede Hopfield, o qual faz a atualização síncrona de pesos, tal que cada nó enxerga também a si próprio. Os dados utilizados na simulação da Hopfield corresponde aos dados da implementação feita com a *backpropagation* (simulação 3). Lembrando que a Hopfield recebe tais dados já devidamente processados, expandidos e normalizados.

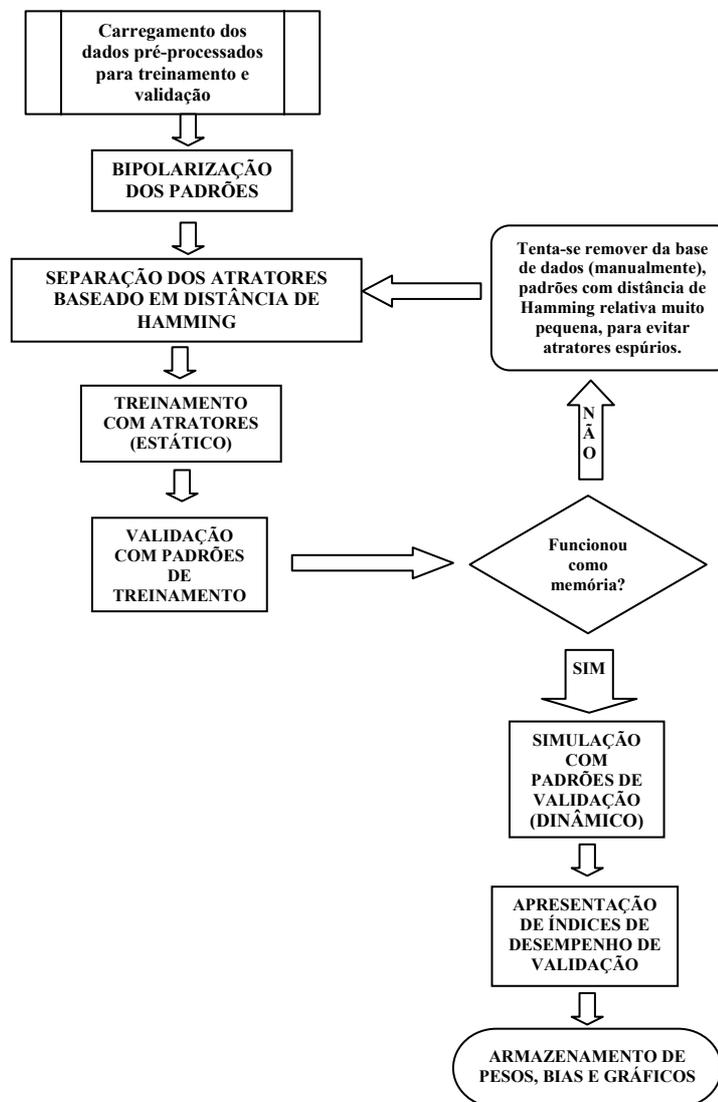


FIGURA 30 – Diagrama do processo de treinamento e validação da Rede Hopfield.

4.12.1.3 – Resultados do Treinamento

A partir de uma escolha dos atratores na base de dados de treinamento foi obtido um vetor com 21 padrões para o treinamento da rede Hopfield. Tais padrões foram escolhidos, aplicando o critério da menor distância de Hamming, de forma que o padrão fosse projetado em posição equidistante em relação aos outros padrões que indicavam a mesma classe.

Foram utilizados 250 bits para codificar as entradas e mais 5 para as saídas, resultando em 255 bits para representar cada padrão.

A rede Hopfield comporta-se como memória: se a entrada for um atrator já presente na rede então, a saída será o próprio atrator. Assim, não será testada a capacidade de armazenamento da rede para os 21 atratores. O que é interessa é a capacidade de reconhecer um padrão pertencente a uma classe que não foi utilizada no treinamento, ou seja, é desejável que a rede seja capaz de generalizar.

4.12.1.4 – Validação

Na FIG. 31 pode-se observar que a rede Hopfield não apresentou um comportamento satisfatório com relação à generalização pois, apesar de ter tido 100% de acerto na fase de treinamento, a validação apresentou um índice de acerto de apenas 13%.

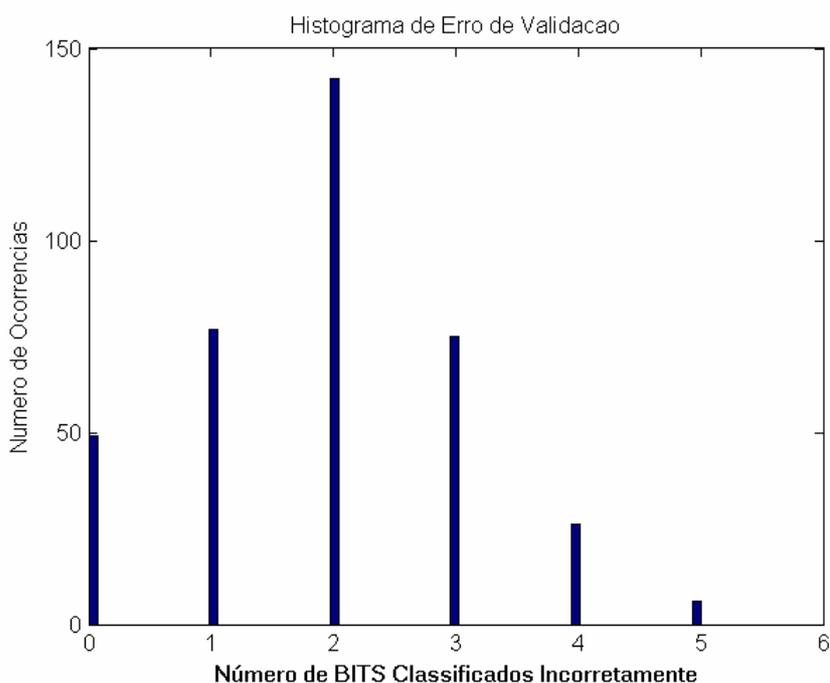


FIGURA 31 – Histograma de Erro na Validação

TABELA 23
Resultados da Rede Hopfield

Parâmetro de Desempenho	Hopfield
• Tempo de treinamento (seg.)	8,4
• Tempo de validação (seg)	322,03
• Dimensão do vetor pesos da camada de entrada	255x255
• Dimensão do vetor pesos da camada de escondida	Não se aplica
• Dimensão do vetor de <i>bias</i>	255x1
• Quantidade total de padrões na base de dados	1126
• Quantidade de padrões de treinamento	21
• Quantidade de padrões de validação	375
• Índice de Acerto no treinamento (em %)	100,0
• Índice de Acerto na validação (em %)	13,07

4.12.1.5 – Conclusões da Tentativa - rede Hopfield

Observando a TAB. 23 pode-se perceber que a Hopfield possui um treinamento estático e bem rápido, mas a recuperação do padrão se faz de maneira dinâmica, o que mostra a fragilidade da rede com relação a esse aspecto.

Semelhante ao caso da BAM, a Hopfield possui um acerto na validação pouco satisfatório. Podem ser levantadas algumas suposições: uma delas pode ser a de que os dados binários característicos do problema apresentam certa dependência linear que não pôde ser removida. Outra explicação seria a de que a Hopfield, assim como a BAM não foi capaz de generalizar o problema, já que se trata de uma memória.

4.13 – Rede de Regressão Generalizada (GRNN)

4.13.1 – Simulação 6

Tendo em vista que o melhor desempenho foi o da rede *backpropagation*, utilizando-se a metodologia de simplificação do problema, tomou-se outro tipo de rede neural para se realizar testes comparativos frente aos resultados da **Tentativa 3**, visto que o desempenho de memórias associativas para o problema em questão não foi satisfatório.

4.13.1.1 – Estrutura da Rede

O MATLAB utiliza as funções de sua biblioteca para montar a estrutura da rede que é dependente do tamanho da base de dados de treinamento, assim, os pesos da camada de entrada, da camada escondida e do *bias* dependem do número de padrões presentes na entrada. O único parâmetro que é possível ser definido na rede é o desvio padrão da função radial presente na camada de entrada.

A GRNN foi simulada e validada a partir do seguinte número de padrões:

- quantidade total de padrões na base de dados: 1126.
 - quantidade de padrões de treinamento: 751.
 - quantidade de padrões de validação: 375.
-

4.13.1.2 – Implementação feita

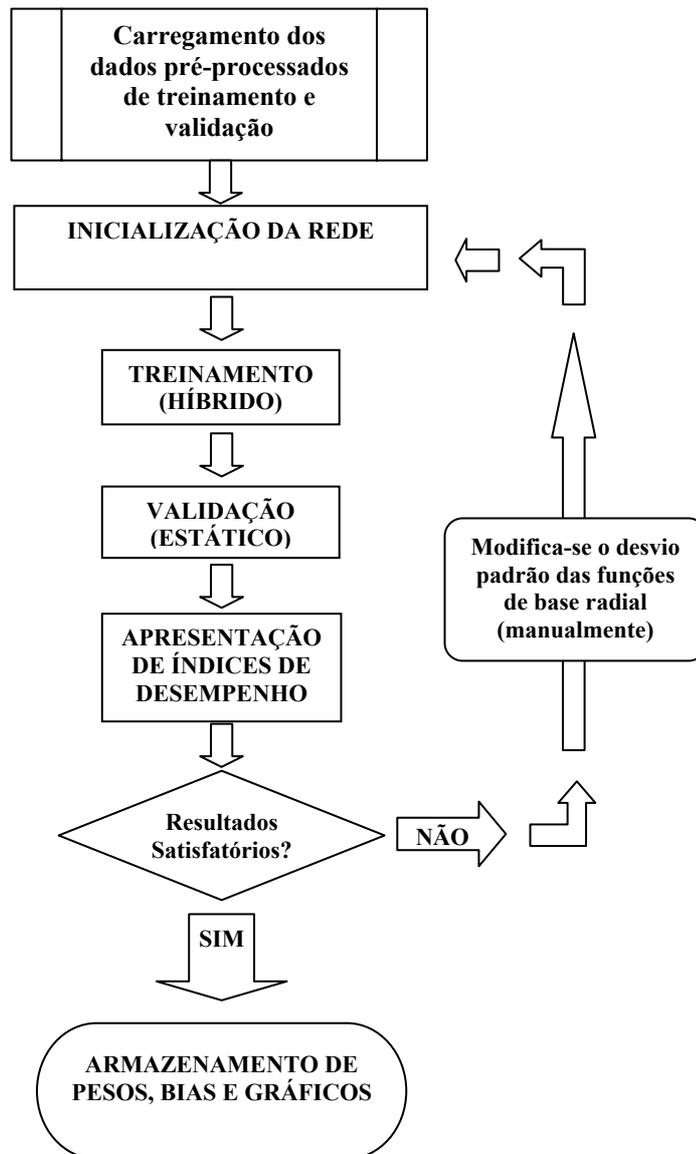


FIGURA 32 – Diagrama do processo de treinamento e validação da rede GRNN.

4.13.1.3 – Resultados do Treinamento

De acordo com o diagrama da FIG. 32 foi gerada uma estrutura para a Rede de Regressão Generalizada. Essa estrutura é definida estaticamente, ou seja, sem treinamento

dinâmico. Devido ao conceito de topologia envolvido, a rede gera um vetor de pesos para cada padrão de entrada. Sendo assim, o tamanho dos vetores dos pesos nas camadas de entrada, camada escondida e vetor de *bias* fica dependente das dimensões do conjunto de treinamento. Tal situação diferencia a GRNN da *backpropagation*, pois nesta, pode-se ajustar o tamanho da matriz de pesos da camada intermediária.

4.13.1.4 – Validação

TABELA 24
Desempenho Estatístico da GRNN durante a validação

Codificação	ACERTOS	ERROS
Bit 1 (curto individual na fase A)	352 (93,87 %)	23
Bit 2 (curto individual na fase B)	366 (97,60 %)	9
Bit 3 (curto individual na fase C)	341 (90,93 %)	34
Bit 4 (Bit 1 da gravidade do curto)	371 (98,93 %)	4
Bit 5 (Bit 2 da gravidade do curto)	357 (95,20 %)	18
Todos os BITS simultaneamente	316 (84,27 %)	59

A TAB. 24 apresenta a análise estatística de desempenho da GRNN para a simulação denominada *Tentativa 3* (existência de curto nas Fases A, B ou C). A análise dos dados demonstra que o índice de acerto na validação foi elevado, ou seja, variou entre 90,93% e

98,93%, considerando os acertos tomados individualmente, por bit. O desempenho da rede quanto às respostas vinculadas à gravidade do curto -bits 4 e 5- na codificação proposta na **Tentativa 3**, também pode ser considerado elevado. A rede também apresentou um bom desempenho quando se considerou o acerto de todos os bits, simultaneamente (84,27%).

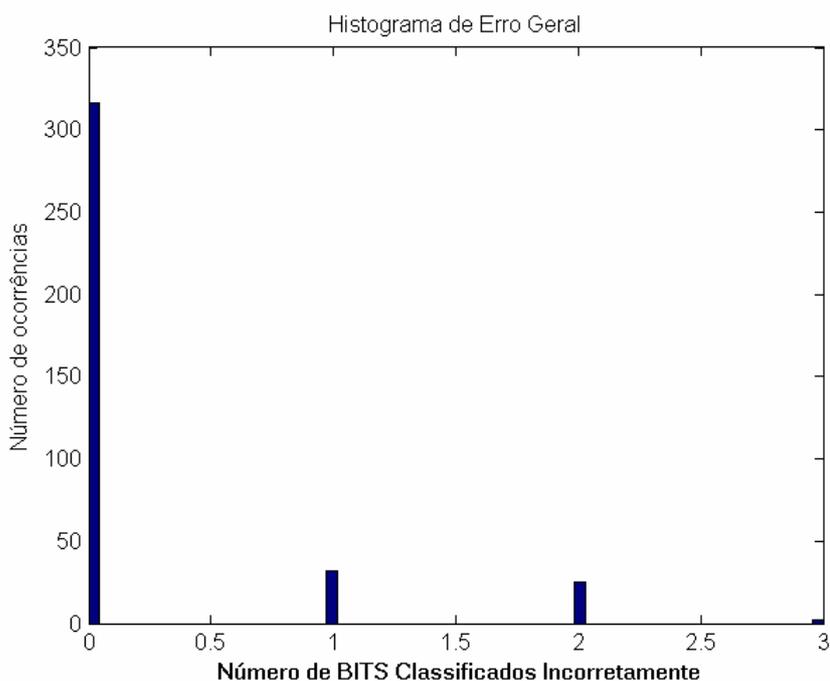


FIGURA 33 – Histograma de erro global, apresentando a quantidade de BITS incorretamente classificados.

A FIG. 33 representa um histograma de erro global na classificação, tomando-se nas abscissas, a quantidade de bits que estavam errados simultaneamente –variando de 1 a 3- durante a fase de validação dos dados e nas ordenadas, a ocorrência de tais situações, ou seja, nenhum bit errado, corresponde a 316 ocorrências conforme a TAB. 24 e FIG. 33. Deve-se ter em mente, no entanto, que a FIG. 33 representa bits que estavam errados, simultaneamente, entre um e três bits.

Pode-se concluir, tanto graficamente quanto pela TAB. 25, que o desempenho da rede foi satisfatório.

TABELA 25
Resultados da Rede GRNN

Parâmetro de Desempenho	GRNN
• Tempo de treinamento (seg.)	6,15
• Dimensão do vetor pesos da camada de entrada	751x5
• Dimensão do vetor pesos da camada de escondida	5x751
• Dimensão do vetor de <i>bias</i>	751x1
• Quantidade total de padrões na base de dados	1126
• Quantidade de padrões de treinamento	751
• Quantidade de padrões de validação	375
• Índice de Acerto no treinamento (em %)	90,27
• Índice de Acerto na validação (em %)	84,26

4.13.1.5 – Conclusões da Tentativa

A rede conseguiu resolver o problema satisfatoriamente. Um dos pontos negativos é justamente o tamanho da estrutura resultante da rede, já que tal estrutura é determinada pelo número de padrões de treinamento (vetor 751 linhas x 5 colunas). O comportamento da rede assemelha-se a uma memória. Apresenta peculiaridades como o fato de trabalhar diretamente com números reais e não com valores binarizados. Porém diferencia-se das memórias por possuir uma camada de saída linear que incorpora à rede a capacidade de aproximador universal de funções, já que a parte não-linear fica sendo responsabilidade da camada das funções radiais.

CAPÍTULO 5

CONCLUSÕES FINAIS E PROPOSTAS DE CONTINUIDADE

5.1 Introdução

Este capítulo é dedicado à avaliação dos resultados obtidos neste trabalho, estando apresentadas as conclusões gerais e propostas de continuidade.

Esta dissertação apresenta várias técnicas e metodologias usadas na concepção de sistemas de diagnóstico de faltas em sistemas físicos, destacando-se o uso de redes neurais artificiais como ferramenta poderosa, capaz de conferir uma apreciável melhoria no gerenciamento, aquisição e processamento de dados, tornando o diagnóstico mais rápido e efetivo, sendo capazes de realizar todas as etapas do processo supervisorio se forem treinadas para isto.

5.2 Resultados Comparativos de Desempenho das Redes Estudadas

A TAB. 26 apresenta um comparativo entre as várias tentativas realizadas e algumas observações em relação às simulações feitas. Também é apresentado um resumo das principais características demonstradas pelas diferentes topologias de redes.

TABELA 26
Resultados Comparativos

REDE	Simulações	Adequação ao problema	Obs:
BACKPROPAGATION	1	Ruim	Problema mal definido.
BACKPROPAGATION	2	Médio	Problema melhor definido, mas ainda complexo.
BACKPROPAGATION	3	Excelente (mais de 80% de acerto)	Simplificação do problema e acerto satisfatório com uma estrutura simples.
BAM	3	Ruim	Desempenho excelente como memória, mas sem generalização.
HOPFIELD	3	Ruim	Desempenho bom como memória, mas sem generalização e recuperação de padrões muito lenta.
GRNN	3	Excelente (mais de 80% de acerto)	Desempenho excelente em troca de uma estrutura complexa e com dados redundantes.

5.3 Conclusões Finais

Apesar de Wasserman (1988) afirmar que tanto a Hopfield como a BAM são capazes de recuperar padrões de entrada contaminados por ruídos, isto não pôde ser observado nas simulações realizadas, com resultados satisfatórios. Tal problema pode ter ocorrido devido aos seguintes fatores:

- Tamanho do conjunto de treinamento ser considerado pequeno;
- Presença de ruídos e/ou informação insuficiente nos dados de entrada;
- A implementação que o MATLAB permite fazer, faz com que as dimensões das matrizes variem exponencialmente em função do número de dados que se deseja memorizar. Isso acarretou erros numéricos, que causaram a degeneração da memória, para um conjunto relativamente grande de padrões.

Observando a TAB. 26, concluímos que as redes *feedforward* apresentam melhores resultados, quando comparadas às memórias. Dentre as *feedforward*, a MLP demonstrou ser a mais adequada ao problema de detecção de faltas no estator de motores de indução trifásicos, pelo fato de, com uma estrutura mais simples, ter sido obtido o melhor desempenho (na simulação 3).

O uso das Redes Neurais Artificiais tipo BAM e Hopfield, para detecção de defeitos em motores de indução trifásicos, possui uma limitação quanto à capacidade de generalização e de armazenar dados. Poder-se-á, no entanto, estender sua aplicação a outros sistemas de diagnóstico de faltas em outras máquinas elétricas como geradores síncronos e transformadores, tomando-se os devidos cuidados na obtenção das massas de dados.

Ficou demonstrado que a rede Perceptron Multicamada, com algoritmo *backpropagation*, apresenta-se como uma ferramenta que pode ser utilizada na manutenção

preditiva, para identificar faltas incipientes, a partir de uma base de dados adequadamente obtida.

As redes desenvolvidas neste trabalho podem tornar-se confiáveis e aplicáveis a motores também na condição sob carga, para detecção de defeitos elétricos. Tal aplicação estará condicionada a um incremento da quantidade de dados de treinamento, para abranger maior quantidade de situações que caracterizem defeitos.

5.4 Metodologia para utilização prática da rede neural

As redes poderão ser treinadas para funcionamento junto a outras máquinas elétricas, com estrutura física semelhante em termos de funcionamento sob condições normais e de defeito, desde que a massa de dados operacionais obtida tenha a mesma formatação da utilizada na aplicação aqui apresentada ou os programas sejam adaptados para novo leiaute de base de dados.

Etapas para geração de nova base de treinamento, sem utilizar placa de aquisição de dados:

- Obter os dados nominais da máquina em termos de tensão, corrente e potência;
 - Aplicar as tensões de alimentação, variando entre zero e a tensão nominal;
 - Anotar os valores das tensões de linha e das correntes das três fases;
 - Fazer o pré-processamento dos dados obtidos;
 - Submeter os valores encontrados às Redes Neurais Artificiais com diferentes topologias.
-

- Fazer treinamento e análise de resultados;
- Observar os resultados que indicarão a(s) fase(s) onde ocorreu(ram) o(s) curto(s), baseado no treinamento que recebeu.

Portanto, conclui-se que, a rede Perceptron Multicamada, com algoritmo *backpropagation*, desenvolvida neste trabalho apresentou o melhor desempenho, estando apta a identificar faltas no estator, conforme lhe sejam apresentados dados obtidos experimentalmente ou via simulação de defeitos. A análise dos resultados obtidos mostrou também que este tipo de rede é aplicável a processos de diagnóstico com reconhecimento de padrões.

5.5 Propostas de Continuidade

O trabalho aqui realizado pode ser continuado através de:

- ⇒ Montagem de uma base de dados que contemple uma maior variedade de faltas no motor de indução, sejam faltas no estator ou no rotor. Promover novo treinamento das redes, em função da nova base de dados obtida. Submeter a rede a testes visando detectar seu comportamento frente aos testes de capacidade de generalização e aplicação a motores com outras características, aproveitando-se do fato de que trabalha com valores “pu”;
-

-
- ⇒ As redes poderão ser utilizadas, se devidamente treinadas, para funcionamento junto a outras máquinas elétricas como geradores e motores síncronos e transformadores;
 - ⇒ Realizar adequações de forma a permitir que as redes trabalhem conectadas a uma placa de aquisição de dados, tornando-se uma ferramenta on-line de diagnóstico de faltas;
 - ⇒ Adaptar e integrar a rede neural Perceptron Multicamada, com algoritmo *backpropagation* a sistemas nebulosos para detecção de defeitos, caracterizando uma rede neurofuzzy, através do qual seja feita a classificação dos padrões de falta em sistemas de diagnóstico em equipamentos elétricos.
-

REFERÊNCIAS BIBLIOGRÁFICAS

ALBRIGHT, D. R. Interturn short-circuit detector for turbine-generator rotor windings. **IEEE Transactions on Power Apparatus and Systems**, v. PAS-90, n.2, p.478-483, March/April, 1971.

ARAÚJO, Filipe João Boavida de Mendonça Machado. **Aplicação das Redes de Hopfield no Encaminhamento em Redes de Dados**. 1999. 201f. Dissertação (Mestrado em Engenharia Informática) - Faculdade de Ciências e Tecnologia Universidade de Coimbra.

ASHEN, R. A; STEELE, M. E. Motor Condition Monitoring Methods , **Procedures of Conference On Drives, Motors and Controls**, Harrogate, North Yorkshire, UK, p. 163-171, Oct. 1983.

BENGIO, Y.; SIMARD, P.; FRASCONI, P. Learning long-term dependencies with gradient descent is difficult. **IEEE Transactions on Neural Networks**, n.5, p.157-166, 1994.

BERNIERI, Andrea; BETTA, Giovanni; LIGUORI, Consolatina. On-line Fault Detection and Diagnosis Obtained by Implementing Neural Algorithms on a Digital Signal Processor. **IEEE Transaction on Instrumentation and Measurement**, v. 45, n.5, p. 894 - 899, Oct. 1996.

BONNETT, Austin H.; SOUKUP, George C. Cause and Analysis of Stator and Rotor Failures in Three-Phase Squirrel-Cage Induction Motors. **IEEE Transactions on Industry Applications**, v. 28, n.4, p.921-937, Jul./Aug. 1992.

BORGARD, Diane E.; OLSSON, Gustaf; LORENZ, Robert D. Accuracy Issues for Parameter Estimation of Field Oriented Induction Machine Drives. **IEEE Transactions on Industry Applications**, v. 31, n. 4 , p.795-801, Jul./Aug. 1995.

CAMERON, J. R.; THOMSON, W. T.; M.I.E.E. e DOW, A B. Vibration and current monitoring for detecting airgap eccentricity in large induction motors. **IEE PROCEEDINGS**, v.133, Pt. B, n.3, p. 155 - 163, May 1986.

CASH, Marcus Alex; HABETLER, Thomas G.; KLIMAN, Gerald B. Insulation Failure Prediction in AC Machines Using Line-Neutral Voltages. **IEEE Transactions on Industry Applications**, v. 34, n.6, p.1234-1239, Nov./Dec. 1998.

CHOW, Mo-Yuen; YEE, Sui Oi. Real Time Application of Artificial Neural Networks for Incipient Fault Detection of Induction Machines. **Proceedings of the 3rd International Conference of Industrial and Engineering Applications of Artificial Intelligence and Expert Systems**, Charleston, South Carolina, July 15-18, p. 1030-1036, 1990.

CHOW, Mo-Yuen ; YEE, Sui Oi. Methodology for on-line incipient fault detection in single-phase squirrel-cage induction motors using artificial neural networks - **IEEE Transactions on Energy Conversion**, v.6, n.3, p.536-545, Sept. 1991.

CHOW, Mo-Yuen ; MANGUM, Peter M.; YEE, Sui Oi. A Neural Network Approach to Real-Time Condition Monitoring of Induction Motors. **IEEE Transactions on Industrial Electronics**, v. 38, n.6, p.448-453, Dec. 1991.

CHOW, Mo-Yuen. **Methodologies of using Neural Network and Fuzzy Logic Technologies for Motor Incipient Fault Detection**. Singapore: World Scientific Publishing Co. Pte. Ltd., 1997. Cap. 1, p.1-11. Cap. 2, p.15-26. Cap. 5, p. 63-78.

CLARK, Paul E.; McSHANE, Ian E. e WAKELEY, Keith. The Diagnosis and Solution of Electrical Machine Problems – Part I: Design Influences. **IEEE Transactions on Industry Applications**, v. IA-23 , n.1, p.28-36, Jan./ Feb. 1987.

FILIPPETTI, F.; FRANCESCHINI, G.; MARTELLI, M.; TASSONI C. An approach to a knowledge representation about induction machine diagnostics in expert systems. **ICEM**, p. 289-296,1988.

FILIPPETTI, F. ; FRANCESCHINI, G.; TASSONI, C. Neural Networks Approach to Electric Machine On-line Diagnostics. **The European Power Electronics Association**, p.213-218, 1993.

FILIPPETTI, F.; FRANCESCHINI, G.; TASSONI, C. Neural Networks Aided On-line Diagnostics of Induction Motor Rotor Faults'. **IEEE Transactions on Industry Applications**, v. 31, n. 4, p.892-899, Jul/Aug 1995.

GONZALEZ, Avelino J.; LOWENFELD, Simon. On-line Diagnosis of Turbine-Generators Using Artificial Intelligence , **IEEE Transactions on Energy Conversion**, v. EC-1,no. 2, Jun 1986, p.68-74.

HARGIS, C.; GAYDON, B. G.; KAMASH, K. The Detection of Rotor Defects in Induction Motors. **Proc. IEE AMDA Conference**, London, England, p.216-220, Sep. 1982.

HAYKIN, S. H. **Neural Networks: A comprehensive foundation**. MacMilliam College Publishing, New York, USA, p. 45-89, 1994.

HERTZ, J. A.; KROGH, A.; PALMER, R. G. **Introduction to the Theory of Neural Computation**. Addison-Wesley Publishing Company, Redwood City, USA, 1991.

HSU, John S. Monitoring of Defects in Induction Motors Through Air-Gap Torque Observation. **IEEE Transactions on Industry Applications**, v. 31, n.5, p.1016-1021, sep./oct. 1995.

ISERMANN, Rolf. Process Fault Detection Based on Modeling and Estimation Methods - A Survey – **Automatica**, v. 20, n. 4, p. 387-404, 1984.

JAIN, Anil K.; MAO, Jianchang; MOHIUDDIN, K. M. Artificial Neural Networks: A Tutorial. [S.l.] **Computer**, p.31-44, Mar. 1996.

JONES, C. R.; TSANG, C. P. On the Convergence of Feed Forward Neural Networks Incorporating Terminal Attractors. **IEEE International Conference on Neural Networks**, p. 253-259, 1993.

KARTALOPOULOS, Stamatios V. **Understanding Neural Networks and Fuzzy Logic - Basic Concepts and Applications**. IEEE PRESS, New York, USA, 1996. 205p.

KERZENBAUM, I.; LANDY C.F. The Existence of Large Inter-Bar Currents in Three Phase Cage Motors with Rotor-bar and End-ring Faults , **IEEE Transaction on Power Apparatus and Systems**, v. PAS-103, n. 7, p. 1854 - 1862, July 1984.

KRAUSE, Paul C.; THOMAS, C. H. Simulation of Simmetrical Induction Machinery – **IEEE Transactions on Power Apparatus and Systems**, v. PAS-84, n. 11, p. 1038-1053, nov. 1965.

KRAUSE, Paul C. **Analysis of Electric Machinery**. United States of America, McGraw-Hill Book Company, 1986, 564 p.

KLIMAN, G. B.; KOEGL, R. A.; STEIN, J.; ENDICOTT, R. D.; MADDEN, M. W. Noninvasive Detection of Broken Rotor Bars in Operating Induction Motors. **IEEE Transactions on Energy Conversion**, Vol. 3, n. 4, p.873 - 879, Dec., 1988.

KLIMAN G. B.; STEIN J. Induction Motor Fault Detection Via Passive Current Monitoring. **Proceedings International Conference on Electric Machines**, Cambridge, Massachusetts, USA. MIT, p.13-17, Aug. 1990.

KOSKO, B. **Neural Networks and Fuzzy Systems**. A dynamical systems approach to machine intelligence, Prentice - Hall International, Inc., USA, p. 55-69, 1992.

KOUROSH Danai; CHIN, Hsinyung. Fault Diagnosis With Proces Uncertainty. **Journal of Dynamic Systems, Measurement, and Control**, v. 113, p. 339-343, Sept. 1991.

KOVÁCS, Z. L. **Redes Neurais Artificiais - Fundamentos e Aplicações**. Edição Acadêmica, São Paulo, Brasil, 1996, 163 p.

LEMAIRE-SEMIL, B.; BOUILLAUT, F.; RAZEK, A. Modelling of vector controlled cage induction motor with FEM. **IEE PROCEEDINGS-B**, v. 138, n. 6, p.297-302, Nov. 1991.

MAIER, Reinhard. Potection of Squirrel-Cage Induction Motor Utilizing Instantaneous Power and Phase Information. **IEEE Transactions on Industry Applications**, v. 28, n.2, p.376-380, March/April 1992.

MATEUS, G. R.; LUNA, H. P. L. *Programação Não-Linear*. V Escola de Computação. UFMG, Belo Horizonte, Brasil, 1986.

MILNE, Robert. Strategies for Diagnosis. **IEEE Transations on Systems, Man, and Cybernetics**, v. SMC-17, no. 3, p. 333-339, May/June 1987.

NATARAJAN, R. Failure Identification of Induction Motors by Sensing Unbalanced Stator Currents. **IEEE Transactions on Energy Conversion**, v. 4, n.4, p. 585-590, Dec., 1989.

NOVOTNY, D. W.; LIPO, T. A. **Dynamics and Control of Induction Motor Drives**. Claredon Press Oxford, 1995. 416p.

OSTOVIC, V. A Simplified Approach to Magnetic Equivalent-Circuit Modeling of Induction Machines. **IEEE Transactions on Industry Applications**, v.24 , n.2, p.1126-1131, March/April, 1988.

OSTOVIC, V. **Dynamics of Saturated Electric Machines**. New York: Springer-Verlag New York Inc., 1989. 445 p.

PENMAN, J.; YIN C.M. The Application of Artificial Neural Networks in Identifying Faults in Induction Machines. **ICEM'92**, Manchester, England, v.3, p.1256-1260, Sept 1992.

PENMAN, J.; YIN, C. M. Feasibility of Using Unsupervised Learning, Artificial Neural Networks For The Condition Monitoring of Electrical Machines. **IEE Proc. - Electr. Power Appl.**, v. 141, n. 6, p.317-322, Nov. 1994.

PROTOPAPAS, C. A.; KAMINARIS, S. D.; MACHIAS, A V.; PAPADIAS, B. C. An Expert System for Fault Repairing and Maintenance of Electric Machines. **IEEE Transactions on Energy Conversion**, v. 5, no. 1, March 1990, p. 79-83.

REASON, John. Pinpoint Induction Motor Faults by Analyzing Load Current. **Power**, p. 87-88, Oct. 1987.

RICH, Elaine; KNIGHT Kevin. **Inteligência Artificial**, 2^a edição, São Paulo, MAKRON Books do Brasil Editora LTDA, Editora McGraw-Hill Ltda, 1993, 722 p.

RILEY, Caryn M.; LIN, Brian K.; HABETLER, Thomas G.; SCHOEN, Randy R. A Method for Sensorless On-line Vibration Monitoring of Induction Machines. **IEEE Transactions on Industry Applications**, v.34 , n.6, p.1240-1245, Nov./ Dec., 1998.

SCHOEN, Randy R.; HABETLER, Thomas G.; KAMRAN, Farrukh; BARTHELD, Robert G. Motor Bearing Damage Detection Using Stator Current Monitoring , **IEEE Transactions on Industry Applications**, v. 31, n.6, p.1274-1279, nov./dec. 1995.

SCHOEN, Randy R.; LIN, Brian K.; HABETLER, Thomas G.; SCHLAG, Jay H.; FARAG, Samir. An Unsupervised, On-Line System for Induction Motor Fault Detection Using Stator

Current Monitoring. **IEEE Transactions on Industry Applications**, v. 31, n. 6, p.1280-1286, Nov./Dec. 1995.

SIYAMBALAPITIYA, D. J. Tilak; MCLAREN, Peter G. Reliability Improvement and Economic Benefits of On-Line Monitoring Systems for Large Induction Machines. **IEEE Transactions on Industry Applications**, v. 26, n. 6, p.1018-1025, Nov./Dec. 1990.

SOHLBERG, Björn. **Supervision and Control for Industrial Processes** – Using Grey Box Models, Predictive Control and Fault Detection Methods , London, England, Springer- Verlag London Limited, 1998, 223 p.

SOOD, Arun K.; FRIEDLANDER, Carl B.; FAHS, Ali Amin. Engine Fault Analysis: Part I – Statistical Methods , **IEEE Transactions on Industrial Electronics**, v.IE-32, n.4, p.294-300, Nov. 1985.

SOOD, Arun K.; FAHS, Ali Amin; HENEIN, Naeim A. Engine Fault Analysis: Part II – Parameter Estimation Approach , **IEEE Transactions on Industrial Electronics**, v.IE-32, n.4, p. 301-307, Nov. 1985.

SORSA, Timo; KOIVO, Heikki N. ; KOIVISTO, Hannu. Neural Networks in Process Fault Diagnosis. **IEEE Transactions on Systems, Man, and Cybernetics**, v.-21, n.4, p.815-825, Jul./Aug. 1991.

STONE, Greg; KAPLER, Joe. Stator Winding Monitoring. **IEEE Industry Applications Magazine**, p.15-20, Sep./Oct. 1998.

SULLIVAN, Charles R.; SANDERS, Seth R. Models for Induction Machines With Magnetic Saturation of the Main Flux Path. **IEEE Transactions on Industry Applications**, v. 31, n. 4, p.907-917, July/Aug. 1995.

TAVNER, P. J.; GAYDON, B. G.; WARD, D. M. Monitoring generators and large motors. **IEE PROCEEDINGS**, v.133, Pt. B, n.3, p. 169 - 180, May 1986.

TAVNER, Peter J.; PENMAN James. **Condition Monitoring of Electrical Machines** , Letchworth, Hertfordshire, England, Research Studies Press Ltd., John Wiley & Sons Inc., 1987, 286 p.

TRUTT, F. C. Steady state analysis of wound-rotor induction machines with simultaneous stator and/or rotor faults. **Electric Machines and Power Systems**, 16, Hemisphere Publishing Corporation, p.35-48,1989.

VAS, Peter. **Parameter Estimation, Condition Monitoring and Diagnosis of Electrical Machines**. New York, Oxford University Press Inc., 1993, 360 p.

WASSERMAN, Philip D. **Neural Computing – Theory and Practice**, New York, Van Nostrand Reinhold, 1989. 230 p.

WILLIAMSON, S.; ABDEL-MAGIED, M.A S. Steady-state Analysis of Double-cage Induction Motors with Rotor-cage Faults - **IEE Proceedings**, v. 134, Pt. B, n. 4, p.199-206, July,1987.

WILLIAMSON, S.; A C.G.I.; D.I.C.; M.I.E.E. e LAITHWAITE, E. R. . Generalised harmonic analysis for the steady-state performance of sinusoidally-excited cage induction motors , **IEEE PROCEEDINGS**, v. 132, Pt. B, n.3, p. 157 - 163, May 1985.

WILLIAMSON, S. e MIRZOIAN, K. Analysis of Cage Induction Motors with stator Winding Faults. **IEEE Transaction on Power Apparatus and Systems**, v. PAS-104,n.7, p. 1838 - 1842, Jul. 1985.

WILLIAMSON, S.; A C.G.I.; D.I.C.; M.I.E.E. e SMITH, A C. Steady-state analysis of 3-phase cage motors with rotor-bar and end-ring faults. **IEEE PROCEEDINGS**, v.129, Pt. B, n.3, p. 93 - 100, May 1982.

YAMAMURA, Sakae. **Spiral Vector Theory of AC Circuits and Machines**, Edição em inglês, New York, Oxford University Press, 1992, 132 p.

REFERÊNCIAS BIBLIOGRÁFICAS UTILIZADAS NA REVISÃO DA DISSERTAÇÃO

BRAGA, Antônio de P. **Redes Neurais Artificiais**. Departamento de Engenharia Eletrônica da Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, 14 set. 2002. Disponível em <<http://www.cpdee.ufmg.br/~apbraga>> Acesso em: 03 dez. 2003.

HAYKIN, SIMON. **Redes Neurais: Princípios e prática**. 2. ed. Trad. Paulo Martins Engel. Porto Alegre: Bookman Companhia Editora, 2001. 900p. cap. 4, p.183-273 e cap. 5, p. 283-343. Original inglês.

MENEZES, Marcio A. F. **Aplicações de Redes Neurais Atratoras no Reconhecimento de Padrões e em Otimização Combinatorial**. 1998. 76f. Dissertação (Mestrado em Física) – Universidade Federal Fluminense.

OGATA, Katsuhiko. **Engenharia de Controle Moderno**, Rio de Janeiro: Editora Prentice/Hall do Brasil Ltda. 1982. Cap. 15, p.810-847.

PARI SOTO, Miguel. **Processamento digital de imagens macroarray para detecção dos resultados de hibridização nos processos de estudos de expressão gênica utilizando redes neurais**. 2003. 139f. Tese (Doutorado em Ciências em Engenharia Civil) – Universidade Federal do Rio de Janeiro.

PINTO, Alan Kardec; XAVIER, Júlio Aquino Nascif. **Manutenção Função Estratégica**. 2. ed. Rio de Janeiro: Qualitymark Editora Ltda, 2003. Cap. 3, p.33-54.

SANTANA, Carlos E. de C.; NASCIMENTO Jr., Cairo L.; SAKANE, Fernando T. Aplicação de redes neurais de Hopfield para detecção e estimação de movimento entre dois quadros de um sinal de vídeo digital. In: SIMPÓSIO BRASILEIRO DE AUTOMAÇÃO INTELIGENTE, 6, 2003, Bauru, São Paulo. **Anais...** [S.l.] [s.n.], 2003. p.337-343.

SHAPIRO, Jonathan. Lecture 5.1 - The Hopfield Model As Auto-associative Memory. **CS6482 - Neural Networks**. Department of Computer Science, University of Manchester, Manchester, 2 fev. 2003. Disponível em <<http://www.cs.man.ac.uk/~jls/CS6482/>> Acesso em: 06 dez. 2003.

THE MATHWORKS INCORPORATION. Matlab. The language of technical Computing versão 6.5.0. 18 jun. 2002. **Neural Network Toolbox User's Guide**. Disponível em <<http://www.mathworks.com>>. Acesso em: 04 dez. 2003.

VELLASCO, Marley Maria B. R. **Inteligência Computacional Aplicada**. Núcleo de Pesquisa de Inteligência Computacional Aplicada da Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 25 out. 2003. Disponível em <<http://www.ica.ele.puc-rio.br/cursos>> Acesso em: 04 dez. 2003.

APÊNDICE A

SISTEMA DESENVOLVIDO

Para possibilitar a reutilização do presente estudo foi desenvolvido um sistema, na plataforma MATLAB, para ser submetido a novos testes com diferentes bases de dados.

Existe apenas um requisito para o funcionamento do sistema. No diretório do sistema existe um subdiretório “\dados”, que deve conter um arquivo chamado “**dados.txt**” que é um arquivo ASCII no seguinte formato:

V	IN	IA	IB	IC	NA	NB	NC	COD1	COD2
75.3	3.0530303	2.6	2.4	3.1	0	0	0	0	0
50	2.520891	2.8	2.7	3.6	0	0	1	0	0
60	2.7681662	2.8	2.5	3.5	0	0	1	0	0
60.2	2.7725426	2.8	3.6	3.4	0	1	1	0	0
50	2.520891	2.9	3.7	3	0	1	0	0	0
60	2.7681662	2.9	3.9	4.1	0	1	1	0	1
70	2.9645161	2.9	2.3	2.9	0	0	0	0	0
70.2	2.968026	2.9	3.8	2.9	0	1	0	0	0
72	2.9989866	2.9	2.4	2.9	0	0	0	0	0

FIGURA 34– Formato do arquivo *batch* de dados.

As colunas são respectivamente: tensão, corrente nominal para a tensão aplicada, corrente na fase A, corrente na fase B, corrente na fase C, bit de curto na fase A, bit de curto na fase B, bit de curto na fase C e mais 2 bits para codificar a gravidade do curto.

Dentro do mesmo diretório “\dados” existe uma rotina “gera_arq” que deve ser executada antes de qualquer treinamento das redes, pois ela gera os arquivos utilizados no treinamento e na validação a partir do arquivo “dados.txt”.

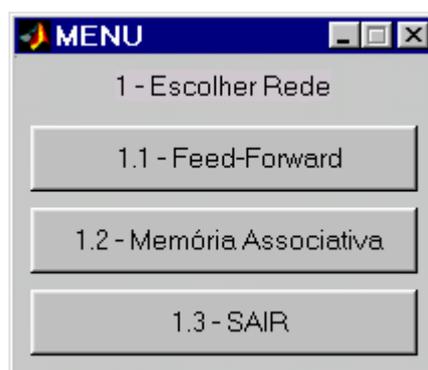


FIGURA 35 – Tela 1 do sistema desenvolvido em MATLAB

A primeira tela, que pode ser visualizada na FIG. 35, apresenta um menu que permite ao usuário selecionar o tipo de filosofia a ser utilizada para resolver o problema.



FIGURA 36 – Tela apresentando as opções de redes *feedforward*.

A segunda tela do sistema (FIG. 36) também é um menu de seleção para que o usuário escolha qual a topologia de redes *feedforward* disponíveis para utilizar no problema.

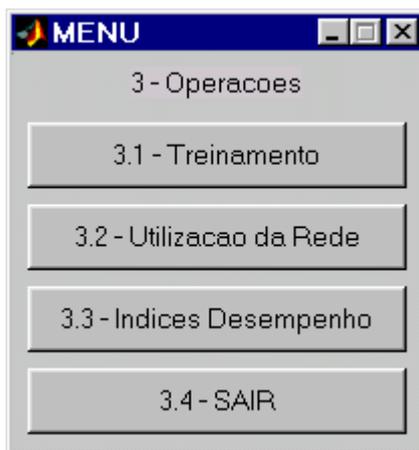


FIGURA 37 – Tela disponível para cada opção de rede, apresentando as operações disponíveis.

A terceira tela (FIG. 37) mostra as operações que podem ser realizadas com a rede escolhida na tela 2. Na opção treinamento ocorre a utilização dos arquivos gerados no diretório “\dados” para realizar-se o treinamento e a validação da rede. A opção 3.2 oferece o diagnóstico da rede para um padrão apresentado a ela. A opção 3.3 fornece os índices de desempenho da rede que são tempo de treinamento, acerto percentual na validação, etc, conforme FIG. 38.

DATA	REDE	TEMPO(seg)	ACERTO TREINAMENTO (%)	ACERTO VALIDACAO (%)
DD-MMM-AAAA	BACKPROPAGATION	165.270	95.78	78.06

FIGURA 38 – Tabela contendo índices de desempenho do treinamento da rede

APÊNDICE B

DADOS OBTIDOS A PARTIR DOS ENSAIOS REALIZADOS NO MOTOR DE INDUÇÃO TRIFÁSICO

B.1. Ensaio a vazio:

a) Medição das resistências do estator e rotor:

$$R_{lcc} : 257\text{m}\Omega \quad (\Omega/\text{fase})$$

$$R_{rcc} : 960\text{m}\Omega \quad (\Omega/\text{fase})$$

Temperatura ambiente: $T_a = 25^\circ\text{C}$

$$R_{lCA} = 1,2 \times R_{lcc} = 308,4 \text{ m}\Omega \quad (19)$$

$$R_{rCA} = 1,2 \times R_{rcc} = 1152 \text{ m}\Omega \quad (20)$$

→ Resistências referidas a 75°C :

$$R_{l(75^\circ\text{C})} = R_l \times \frac{234,5 + 75}{234,5 + T_a} = 367\text{m}\Omega \quad (21)$$

$$R_{r(75^\circ\text{C})} = 1374\text{m}\Omega \quad (22)$$

b) Medições feitas:

TABELA 27
Medições feitas durante o ensaio a vazio

TENSÃO	CORRENTE				Potência
V	A1	A2	A3	Io	Po (W)
220	1,97	1,75	1,98	1,90	220

c) Cálculo das perdas magnéticas:

$$P_{jo} = R_l \times I_o^2 = 367\text{m}\Omega \times 1,982\text{A} = 1,44\text{W}; \quad (23)$$

$$P_{HF} = (P_o/3) - P_{jo} = (220/3) - 1,44 = 71,9\text{W} \quad (24)$$

d) Cálculo da corrente de perdas:

$$I_{HF} = \frac{\sqrt{3} \times P_{HF}}{V} = \frac{\sqrt{3} \times 71,9}{220} = 0,566\text{A} \quad (25)$$

e) Cálculo da corrente de magnetização:

$$I_m = \sqrt{I_o^2 - I_{HF}^2} = \sqrt{1,98^2 - 0,566^2} = 1,89\text{A} \quad (26)$$

f) Cálculo do ramo paralelo do circuito equivalente:

$$g_c = \frac{\sqrt{3} \times I_{HF}}{V} = \frac{\sqrt{3} \times 0,566}{220} = 4,46\text{mS}^1 \quad (27)$$

$$b_m = \frac{\sqrt{3} \times I_m}{V} = \frac{\sqrt{3} \times 1,89}{220} = 14,87\text{mS} \quad (28)$$

B.2. Ensaio de Curto-Circuito:

TABELA 28
Medições feitas durante o ensaio em curto-circuito

TENSÃO	CORRENTE							POTÊNCIA
V	w1	v1	u1	W1	V1	U1	Ie	Pj
53,8	4,0	3,9	3,9	20,0	19,5	19,5	19,7	800

Calculando os valores de Xcc e Rcc:

a) Cálculo da impedância de curto-circuito:

$$Z_{cc} = \frac{V_{cc}}{\sqrt{3} \times I_e} = \frac{53,8}{\sqrt{3} \times 19,7} = 1,58\Omega \quad (29)$$

¹ S → Siemens = $\frac{1}{\Omega}$

b) Resistência Equivalente:

$$R_{cc} = \frac{P_j}{3I_e^2} = \frac{800}{3 \times 19,7^2} = 0,696 \Omega \quad \text{para } T_a : 25^\circ\text{C} \quad (30)$$

c) Reatância de Curto-Circuito:

$$X_{cc} = \sqrt{Z_{cc}^2 - R_{cc}^2} = \sqrt{1,58^2 - 0,696^2} = 1,466\Omega \quad (31)$$

d) Reatância de dispersão - é calculada a partir do tipo de rotor:

$$\text{Para o motor cujo rotor é enrolado: } X_1 = 0,5 X_{cc} = 0,5 \times 1,466 = 0,733 \Omega \quad (32)$$

$$X_2' = 0,5 X_{cc} = 0,5 \times 1,466 = 0,733 \Omega \quad (33)$$

e) Resistência Virtual do estator e rotor:

$$R_{1_{ca(75^\circ\text{C})}} = 1,2 \times R_{1_{cc(75^\circ\text{C})}} \quad (34)$$

$$Kt = \frac{234,5 + 75}{234,5 + t} \quad (35)$$

$$R_{2'_{(75^\circ\text{C})}} = R_{cc} \times Kt - R_{1_{ca(75^\circ\text{C})}} \quad (36)$$

$$R_{2'_{(75^\circ\text{C})}} = 0,773 \times 1,193 - 0,367 = 0,555\Omega \quad (37)$$

B.3. Memória de cálculo dos parâmetros do motor de 4,6 kW

Indutância magnetizante para todo o enrolamento de uma fase:

$$L_{ms} = \mu_0 N_s^2 \frac{r \ell \pi}{g 4} \quad (38)$$

Consideremos: $N_s = 84$ (número de espiras total por fase);

$$r = 64.425 * 10^{-3} \text{ m (raio médio do entreferro);}$$

$$l = \text{comprimento axial do rotor} = 100 * 10^{-3} \text{ m}$$

$$g = 1.3 * 10^{-3} \text{ m (comprimento do entreferro radial)}$$

$$L_{ms} = 51.75 * 10^{-3} \text{ H} \quad (39)$$

Tomando como base a Reatância de dispersão calculada no item 2:

$$X1 = 0,5 X_{cc} = 0,5 \times 1,466 = 0,733 \Omega \quad (40)$$

$$X2' = 0,5 X_{cc} = 0,5 \times 1,466 = 0,733 \Omega \quad (41)$$

$$X1=X2' = 0.733=2*\pi*f*L \quad (42)$$

Assim:

As indutâncias de dispersão do estator e do rotor têm o mesmo valor:

$$L = 1.94 * 10^{-3} \text{ H} \quad (43)$$

Resistência Virtual do estator e rotor:

$$R1_{ca(75^\circ\text{C})} = 367\text{m}\Omega \quad (44)$$

$$R2'_{(75^\circ\text{C})} = 555\text{m}\Omega \quad (45)$$

APÊNDICE C

REDES NEURAIS ARTIFICIAIS

C.1 – Introdução

Este capítulo é dedicado à apresentação das características gerais de funcionamento das redes neurais artificiais, especialmente daquelas redes escolhidas para estudo: (a) rede Perceptron multicamada; (b) rede neural chamada “BAM” - Memória Associativa Bidirecional; (c) rede Hopfield; (d) Rede de Base Radial - RBF.

C.2 – Conceitos gerais

As Redes Neurais Artificiais (RNA) são inspiradas em processos biológicos, buscando realizar, de forma análoga, as mais elementares funções do neurônio humano, como: capacidade de cálculos e representação distribuída, habilidade de aprendizado e generalização, adaptabilidade, tolerância a faltas, reconhecimento de padrões e baixo consumo de energia.

Conforme citado por Kartalopoulos (1996), Warren McCulloch e Walter Pitts publicaram, em 1943, a referência zero sobre a teoria de redes neurais artificiais e foram os primeiros a propor um modelo computacional para o neurônio biológico (ver Kovács (1996) e Kartalopoulos (1996)), conforme o apresentado na FIG. 39, composto por um conjunto de entradas v_1, v_2, \dots, v_n , uma unidade de processamento Σ_i , uma função de ativação não linear f_i e uma saída O_i . Cada entrada j recebe um estímulo v_j que é ponderado pelo peso w_{ij} , que pode ser excitatório ou inibitório. A unidade de processamento coleta os sinais ponderados e os agrega juntamente com o termo *bias*, ou limiar, θ_i . A soma de todas as entradas ponderadas e o limiar é então modulada pela função de ativação f_i que determina o nível de excitação gerado pelo elemento neural. Os neurônios biológicos respondem de forma similar, porém as funções de ativação utilizadas nos elementos neurais não são réplicas das biológicas e são utilizadas de acordo com o tipo de aplicação.

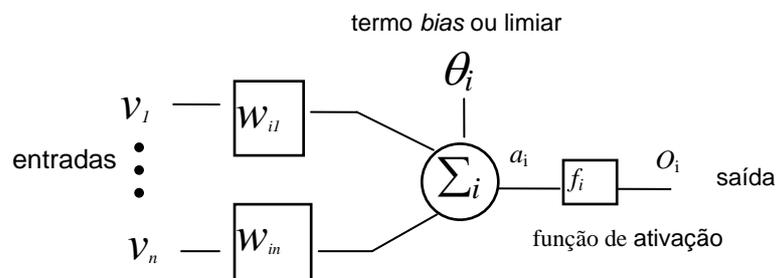


FIGURA 39 – Unidade Neural Básica de McCulloch e Pitts

As funções de ativação mais populares são a sigmoideal, a rampa e a degrau, mostradas na FIG. 40.

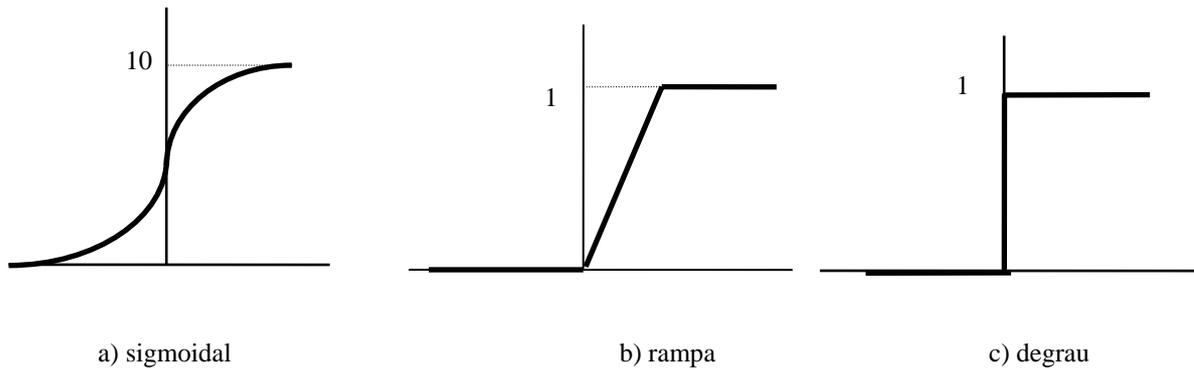


FIGURA 40 – Funções de Ativação Típicas

A saída do modelo básico, mostrado na FIG. 39, pode ser descrita pela seguinte relação:

$$O_i = f_i(a_i), \quad (46)$$

onde

$$a_i = \sum_j w_{ij} v_j - \theta_i. \quad (47)$$

Dentre as funções citadas, a sigmoideal, por ser uma função monotônica, limitada e ter uma derivada fácil de calcular é mais amplamente utilizada em aplicações que envolvem a utilização de redes neurais artificiais. Por exemplo, utilizando-se a função de ativação sigmoideal, a saída do modelo básico mostrado na FIG. 39 será dada por:

$$O_i = f(a_i) = \frac{1}{1 + e^{-a_i}}, \quad (48)$$

que possui uma derivada simples:

$$f'(a_i) = \frac{e^{-a_i}}{(1 + e^{-a_i})^2}. \quad (49)$$

De acordo com Kovács (1996), Rosenblatt cria, no final da década de 50, dando prosseguimento às idéias de McCulloch e Pitts, uma rede com múltiplos elementos neurais dispostos em camadas, a qual chama perceptron. Minsky e Papert mostraram as limitações de

um perceptron simples. Os estudos nesta área permaneceram parados por quase 20 anos, até que em 1980, o interesse pelas redes neurais se renovou. Os maiores favorecedores deste ressurgimento foram a abordagem de energia de Hopfield em 1982 e o algoritmo de aprendizado de retropropagação para redes *feedforward* multicamada, proposto primeiro por Werbos, reinventado várias vezes e popularizado, anos mais tarde, por Rumelhart em 1986.

Assim surgiram as redes neurais artificiais, sistemas compostos por conjuntos de elementos neurais interligados capazes de fazer classificações e de reconhecer padrões geométricos e abstratos.

As RNA's simulam o comportamento de uma rede de neurônios biológicos, através de modelos computacionais que realizam a generalização dos sinais de entrada a partir de um modelo de classificação obtido durante a fase chamada de treinamento.

As principais aplicações de redes neurais artificiais são na resolução de complexos problemas como:

- Classificação de padrões: como em aplicações onde são necessários reconhecimento de caracteres, de voz, inspeção de placas de circuitos impressos, etc.
 - Agrupamento / categorização: não é necessário treinar a rede com classes de dados conhecidas; o algoritmo de agrupamento explora a similaridade entre padrões e os agrupa em conjuntos diferentes, como em aplicações de compressão de dados e análise exploratória de dados.
 - Aproximação de funções: consiste em encontrar uma função que represente um determinado comportamento em problemas de modelagem nas áreas científica e de engenharia.
 - Prognóstico / previsão: dado um conjunto de amostragens em uma certa seqüência no tempo, realizar a previsão de amostragens em tempos futuros. Este tipo de atividade tem
-

um significativo impacto sobre a tomada de decisões em negócios, ciência e engenharia.

Algumas aplicações são: previsão de estoques e de meteorologia.

- Otimização: consiste em encontrar a solução que satisfaça um conjunto de restrições tal que uma função objetiva seja maximizada ou minimizada. Este tipo de técnica tem larga aplicação em matemática, estatística, engenharia, ciências, medicina e economia.
- Memória de conteúdo endereçável: a memória associativa pode ser acessada não pelo endereço, mas pelo conteúdo total ou parcial. A vantagem é obter o acesso mesmo que haja um erro no endereçamento. Um campo de atuação é na construção de bancos de dados com informação multimídia.
- Controle: em modelos de referência adaptativa, o objetivo é gerar uma entrada de controle tal que o sistema forneça uma saída conforme a determinada por um modelo de referência. Um tipo de aplicação é em sistemas dinâmicos para controle de velocidade.

A rede neural artificial possui, como componentes básicos: (a) um conjunto de elementos neurais, (b) uma arquitetura que define o número de elementos e a forma com se interconectam, e (c) propriedades funcionais que definem como a rede aprende, recorda, associa, compara e classifica. A FIG. 41 ilustra a topologia de uma rede neural artificial multicamada. Os elementos que recebem diretamente as entradas da rede constituem o que se chama de camada de entrada. Os elementos que recebem como entradas, as saídas daqueles da camada de entrada, constituem a segunda camada e assim sucessivamente até a camada final, denominada camada de saída. As camadas intermediárias (internas) são geralmente referidas como camadas ocultas. Apesar de não representado, cada elemento neural da FIG. 41 possui aquela estrutura apresentada na FIG. 39.

C.3 – Aprendizagem de uma Rede Neural Artificial

Sistemas biológicos possuem a propriedade de **aprender** uma função, sendo que um dos métodos de treinamento mais intuitivos é aquele feito através de exemplos.

A propriedade mais importante de uma rede neural artificial é, portanto, aprender a partir do ambiente que lhe é apresentado, melhorando seu desempenho, a partir do processo iterativo de ajustes dos pesos e limiares que é o treinamento. O aprendizado ocorre quando a rede neural consegue dar uma solução generalizada para um determinado problema.

Denomina-se **algoritmo de treinamento** ao conjunto de regras que orientam a solução de um problema. A rede neural artificial passa a conhecer o problema, através da apresentação de padrões de entradas/saídas considerados como exemplos representativos do funcionamento do sistema. Algoritmos de treinamento utilizam regras de treinamento que controlam os processos de atualização dos pesos. No entanto, diferentes arquiteturas de redes requerem adequados algoritmos de aprendizado, que, no entanto, podem fazer uso de quatro regras básicas de aprendizado, conforme Jain, Mao e Mohiuddin (1996): correção de erros, regra de Hebb, regra de Boltzmann e aprendizado competitivo.

Durante o processo de treinamento, a rede ajusta os pesos entre elementos e aqueles que ponderam as entradas unitárias dos limiares. Este processo é feito continuamente, em resposta aos estímulos de entrada, produzindo o que Haykin (1994), chama de **aprendizagem**. Quando a saída estiver dentro dos limites estabelecidos como toleráveis, a fase de treinamento estará terminada, tendo a rede adquirido **conhecimento**.

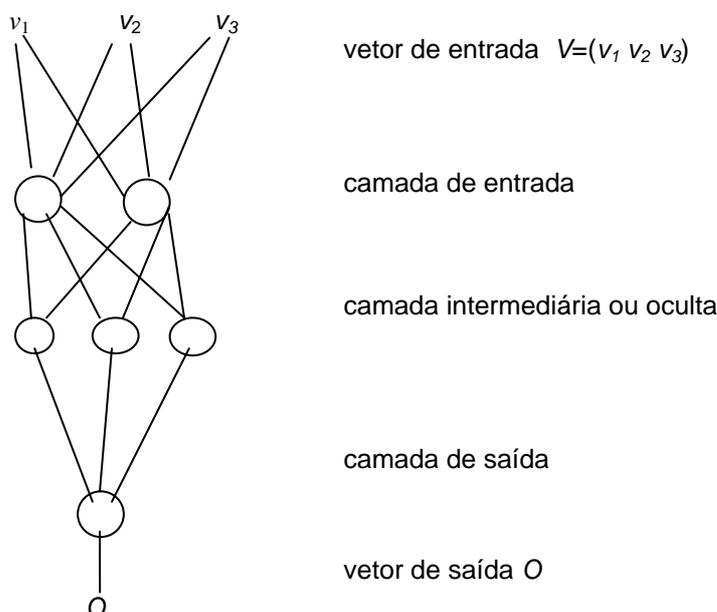


FIGURA 41 – Estrutura de uma Rede Neural Artificial Multicamadas

As redes neurais artificiais podem aprender por memorização, por analogias, através de exemplos, por exploração e também por descoberta, porém, nem todos os processos de aprendizado possuem o mesmo grau de eficiência. Por isto, é necessário escolher dentre as diferentes técnicas de aprendizado disponíveis, aquela que melhor se adapte às características de cada rede e à natureza do problema em consideração.

Algumas filosofias de aprendizagem são:

- **Aprendizagem supervisionada**, processo realizado mediante sucessivas apresentações de padrões de entrada-saída conhecidos, requerendo um “professor” ou “supervisor”, responsável pelo processo de minimização da diferença entre resposta da rede e a resposta esperada, determina ainda os ajustes que devem ser feitos para que o erro seja minimizado;
- **Aprendizagem não supervisionada**: não requer um “professor”, nem existe uma saída esperada. Pelo treinamento, a rede detecta características estatísticas nos dados de entrada

e cria diretrizes para a formação de agrupamentos que constituirão diferentes classes. Ao receber um estímulo, a rede indica a classe em que se encaixa tal estímulo. Se a rede não for capaz de encontrar a classe para o estímulo recebido, uma nova classe será gerada, conforme as diretrizes para seleção de características de agrupamentos, incluídas no projeto da rede.

Wasserman (1989) levantou uma importante questão: o processo supervisionado é inaceitável em termos de analogia com sistemas biológicos, já que o cérebro não compara respostas reais com respostas desejadas. Os treinamentos não supervisionados são mais parecidos com os processos de aprendizado em sistemas biológicos.

- **Aprendizagem por competição:** os nós de saída competem para ser o único a disparar, a partir de um monitoramento centralizado. O nível de ativação maior define o nó vencedor, o qual terá o vetor de pesos atualizado.
- **Aprendizagem por reforço:** esta modalidade permite avaliação apenas qualitativa das respostas ou ações: não há avaliação segundo o critério de correto ou errado, mas, por exemplo, como boas ou más, assim, a realimentação em relação ao meio é avaliativa, não instrutiva. Diferencia-se do aprendizado supervisionado pela inexistência de exemplos e pelo fato de que o ambiente recebe a informação da ação correta, imediatamente.

Há diversos algoritmos de treinamento disponíveis e o método e a velocidade com que os pesos são ajustados depende do tipo utilizado. Os algoritmos de treinamento podem ser vistos como algoritmos de otimização que promovem a minimização da função erro, definida em relação aos valores esperados na saída, obedecendo aos limites de tolerância pré-estabelecidos.

Redes neurais artificiais podem ser agrupadas em duas categorias principais: (a) redes *feedforward* nas quais não há laços fazendo interconexões entre saídas e entradas; (b) redes

recorrentes (*feedback*), nas quais ocorrem conexões entre as saídas e entradas. As redes recorrentes são consideradas dinâmicas porque quando novos padrões de entrada são apresentados à rede, as saídas são recalculadas e, através do laço de realimentação, as entradas são modificadas o que conduz a rede a um novo estado. São consideradas como pertencentes à família de redes *feedforward*: o Perceptron camada simples, o modelo multicamadas e as redes chamadas RBF – Função de Base Radial.

Como pertencentes às redes recorrentes, podemos citar: redes competitivas, redes Mapas auto-organizados de Kohonen, redes Hopfield e redes ART – modelos de redes baseadas na Teoria de Ressonância Adaptativa de Grossberg e Carpenter.

A arquitetura de uma rede é definida pela forma como seus neurônios são conectados. Diferentes formas de conexão entre neurônios produzem diferentes comportamentos para as redes neurais artificiais.

C.4 – Arquitetura: Perceptron multicamadas

Diferentes arquiteturas de redes requerem adequados algoritmos de aprendizado, com diferentes procedimentos de atualização dos pesos e limiares, que são definidos pelas regras de aprendizado. As regras básicas de aprendizado são: por correção de erros, por Boltzmann, Hebbiana e por aprendizado competitivo.

A família de redes neurais *feedforward* mais comum é a chamada Perceptron multicamadas, cujos neurônios são organizados em camadas e possuindo apenas conexões unidirecionais entre eles. As redes Perceptron Multicamadas (MLP) são formadas por perceptrons simples organizados em estrutura hierárquica, que podem ter uma ou várias

camadas intermediárias entre as camadas de entrada e saída. Não existe um método formal para se definir o número adequado de camadas e de nós por camadas. Objetivando encontrar a melhor solução para um problema, normalmente, utiliza-se o método de tentativa e erro.

C.4.1 – Algoritmo de treinamento *backpropagation*

Backpropagation é um método de treinamento que é classificado como pertencente à categoria de aprendizagem supervisionada. Sua aplicação às redes Perceptron multicamadas tem tornado estas redes neurais muito populares entre pesquisadores e usuários de todo o mundo.

Este algoritmo tem sido amplamente utilizado, sendo, conceitualmente, uma generalização do algoritmo de aprendizado de Widrow e Hoff - “*Least Mean Square*” (LMS), também conhecido como Regra Delta. Ver Kosco (1992), Kovács (1996) e Kartalopoulos (1996).

O desempenho da rede é medido por uma função erro que considera, para cada um dos diversos padrões “ p ” de entrada, o quadrado da diferença entre o valor esperado e a respectiva saída calculada. Em outras palavras, o erro é o somatório dos erros quadráticos, definido pela seguinte expressão:

$$E = \sum_{i,p} E_i^p = \frac{1}{2} \sum_{i,p} (d_i^p - O_i^p)^2, \quad (50)$$

onde E_i^p representa o erro no i -ésimo elemento neural, para o p -ésimo padrão de entrada, d_i^p é a saída esperada no i -ésimo elemento neural, para o p -ésimo padrão de entrada e O_i^p é a saída produzida, definida como:

$$O_i^p = f(a_i^p) = f\left(\sum_j w_{ij} v_j^p - \theta_i\right), \quad (51)$$

onde v_j^p é a j -ésima componente do padrão de entrada V^p .

O algoritmo *backpropagation* ajusta os pesos das unidades das camadas intermediárias a partir dos erros das unidades da camada de saída. Para iniciar este processo de aprendizado é necessário estabelecer: (a) um conjunto de padrões de entrada e respectivas saídas esperadas; (b) a taxa de aprendizado γ , (c) um critério de parada, com o qual se dará o treinamento como concluído e (d) uma função de transferência. A aplicação do algoritmo *backpropagation* envolve um deslocamento para frente, através da rede, para calcular as saídas de cada elemento neural da camada de saída e o seu respectivo erro. A segunda fase envolve um deslocamento no sentido contrário, durante o qual o sinal de erro é passado para cada elemento e as mudanças nos pesos são efetuadas. A rede é treinada através da apresentação de todos os dados de treinamento repetidas vezes. Em cada iteração, a diferença entre o valor da saída e o valor esperado, determina a correção que deverá ser feita nos valores dos pesos e dos limiares. Os valores dos pesos são ajustados após cada etapa de treinamento, até que a função erro seja reduzida a um valor aceitável.

O algoritmo *backpropagation* atua nos pesos sinápticos, minimizando a função erro. Existem muitas variações do algoritmo *backpropagation*. A implementação mais simples, chamada “gradiente descendente”, atualiza os pesos e limiares na direção em que a função decai mais rapidamente – o negativo do gradiente. Neste método, os valores dos pesos são modificados proporcionalmente ao oposto da derivada do erro, de acordo com a seguinte expressão:

$$\Delta w_{ij} = -\gamma \frac{\partial E_i^p}{\partial w_{ij}}, \quad (52)$$

Onde γ , denominado taxa de aprendizado, controla quão grande deve ser o “passo” a ser dado. Para maiores detalhes sobre o método do gradiente descendente, recomenda-se, por exemplo, o texto de Mateus & Luna (1986).

Definindo-se $d_i^{p,k}$ como a saída esperada na i -ésima unidade da k -ésima camada, quando o p -ésimo padrão é apresentado à rede e,

$$O_i^{p,k} = f(a_i^{p,k}), \quad (53)$$

como a saída real na unidade, onde $a_i^{p,k} = \sum_j w_{ij}^k O_j^{p,k-1} - \theta_i$ e $O_i^{p,0} = v_i^p$, a função erro, na camada k , poderá ser escrita como:

$$E^k = \frac{1}{2} \sum_{i,p} (d_i^{p,k} - O_i^{p,k})^2. \quad (54)$$

A correção nos pesos na camada de **saída K** é dada pela aplicação da regra da cadeia:

$$\Delta w_{ij}^K = -\gamma \frac{\partial E^K}{\partial w_{ij}^K} = -\gamma \sum_p \frac{\partial E^K}{\partial O_i^{p,K}} \frac{\partial O_i^{p,K}}{\partial a_i^{p,K}} \frac{\partial a_i^{p,K}}{\partial w_{ij}^K} = \gamma \sum_p \delta_i^{p,K} O_j^{p,K-1}, \quad (55)$$

Definindo-se:

$$\delta_i^{p,K} = f'(a_i^{p,K}) (d_i^{p,K} - O_i^{p,K}), \quad (56)$$

Similarmente, pode ser mostrado que nas camadas **intermediárias** a correção é dada por:

$$\Delta w_{ij}^k = -\gamma \frac{\partial E^k}{\partial w_{ij}^k} = \gamma \sum_p \delta_i^{p,k} O_j^{p,k-1}, \quad (57)$$

Definindo-se:

$$\delta_i^{p,k} = f'(a_i^{p,k}) \sum_j w_{ij}^{k+1} \delta_j^{p,k+1}. \quad (58)$$

A seguir, é apresentado em forma algorítmica, o aprendizado *backpropagation*:

algoritmo

inicialize w_{ij}^k com números aleatórios “pequenos”

enquanto $E \geq \varepsilon$ **faça**

para cada padrão $p=1,2,\dots,P$ **faça**

 /* inicialize entradas */

$$O_i^{p,0} \leftarrow v_i^p \quad \forall i=1,2,\dots,I$$

 /* propagar sinal à frente */

para cada camada $k=1,2,\dots,K$ **faça**

$$O_i^{p,k} \leftarrow f\left(\sum_j w_{ij}^k O_j^{p,k-1} - \theta_i\right) \quad \forall i=1,2,\dots,I$$

fim para

 /* calcular delta da camada de saída K */

$$\delta_i^{p,K} \leftarrow f'\left(\sum_j w_{ij}^K v_j^{p,K-1} - \theta_i\right) (d_i^{p,K} - O_i^{p,K}) \quad \forall i=1,2,\dots,I$$

 /* calcular deltas para camadas intermediárias */

para cada camada $k=K-1,\dots,1$ **faça**

$$\delta_i^{p,k} \leftarrow f'\left(\sum_j w_{ij}^k O_j^{p,k-1} - \theta_i\right) \sum_j w_{ij}^{p,k+1} \delta_j^{p,k+1} \quad \forall i=1,2,\dots,I$$

fim para

 /* atualizar pesos */

para cada camada $k=1,2,\dots,K$ **faça**

$$\Delta w_{ij}^k \leftarrow \gamma \sum_p \delta_i^{p,k} O_j^{p,k-1}, \quad \forall i,j$$

$$w_{ij}^k \leftarrow w_{ij}^k + \Delta w_{ij}^k, \quad \forall i,j$$

fim para

fim para

fim enquanto

fim algoritmo

O algoritmo *backpropagation* tem sido utilizado com sucesso em muitas aplicações práticas. Entretanto, como as superfícies de erro não lineares são geralmente irregulares e com vários mínimos locais, ele apresenta alguns problemas, tais como taxa de convergência lenta e propensão a ficar “preso” em um mínimo local, sem conseguir chegar ao mínimo global.

Um outro problema que pode surgir é a chamada “paralisia de treinamento”, quando se utiliza função de ativação sigmoideal, equação (48). Esta paralisia, encontrada durante a fase de treinamento, caracteriza-se por altos valores $a_i^{p,k}$, causados por valores de pesos também

elevados, o que conduz a um gradiente muito pequeno. Na região de gradiente pequeno, a correção dos pesos, que é proporcional ao gradiente, fica paralisada. Algumas técnicas conhecidas para resolver este problema são: a limitação dos valores dos pesos a um intervalo pré-estabelecido ou a redução dos valores iniciais que lhes são atribuídos, bem como aos limiares.

A aplicação da técnica do gradiente descendente com ou sem o momentum, ainda é muito lenta para aplicação em problemas reais. Assim, para aumentar a velocidade de convergência, foram desenvolvidas duas categorias de algoritmos: a chamada categoria heurística e a outra categoria de algoritmos que utiliza técnicas de otimização numérica para aumentar em até uma centena de vezes a velocidade de convergência em relação ao método *backpropagation* padrão.

São exemplos de métodos pertencentes à categoria heurística: a técnica do *momentum*, a variação da taxa de aprendizado durante o processo de treinamento e o método chamado *resilient backpropagation*.

Quanto à categoria de algoritmos que utilizam técnicas de otimização numérica para acelerar o treinamento das redes neurais, podemos citar 3 técnicas: gradiente conjugado, métodos quasi-Newton e o método de Levenberg-Marquardt.

C.4.1.1 – Técnicas do *momentum* e medida da entropia do erro

A técnica do *momentum* consiste em incorporar o termo - μ - ao gradiente descendente, o que permite à rede responder não só aos pequenos gradientes como também, às variações da superfície de erro.

O método *momentum* (Hertz *et al.*, 1991 e Penman & Yin, 1992), é usado para acelerar a convergência da rede e conferir estabilidade. Nesta técnica, a atualização dos pesos é uma ponderação entre o passo atual e os calculados anteriormente. Consiste em aplicar uma constante μ sobre a mudança no vetor de pesos da última iteração, somando-se a esta, a correção de direção da iteração corrente. Se a última mudança de pesos foi em uma direção particular, o termo *momentum* faz com que a próxima mudança de peso seja mais ou menos na mesma direção que a anterior. Dependendo da situação este pequeno desvio pode ser suficiente para evitar que a rede caia em um mínimo local.

A correção, incorporando o *momentum*, passa a ser dada pela expressão:

$$\Delta w_{ij}^k(t+1) = \gamma \frac{\partial E^k}{\partial w_{ij}^k} + \mu \Delta w_{ij}^k(t). \quad (59)$$

Teoricamente o *momentum* possibilita utilizar valores maiores para a taxa de aprendizado, porque se constitui em um filtro passa-baixa. Desta forma, a velocidade de convergência é aumentada. O negativo do gradiente dá a direção para otimização, mas não a magnitude do passo a ser dado. Sendo normal que vários passos na direção do gradiente descendente tenham que ser dados para encontrar o mínimo da função. No ponto mínimo, os valores dos elementos do vetor gradiente serão todos iguais a zero. A aplicação do gradiente descendente pode terminar em qualquer tipo de ponto estacionário: a solução ou um ponto de sela, o que exige a aplicação de outros métodos para encontrar o mínimo global.

A taxa de convergência do processo é fortemente dependente dos valores da taxa de aprendizado “ γ ”. Valores muito grandes desta taxa levam a oscilações e instabilidade, enquanto valores muito pequenos podem reduzir drasticamente a taxa de convergência.

Jones e Tsang (1993) apresentam outra variação desse algoritmo que consiste em escolher uma função quadrática e usar a **medida da entropia relativa**, determinando uma nova função erro:

$$E^k = \sum_{i,p} \left[d_i^{p,k} \log\left(\frac{d_i^{p,k}}{O_i^{p,k}}\right) + (1 - d_i^{p,k}) \log\left(\frac{1 - d_i^{p,k}}{1 - O_i^{p,k}}\right) \right]. \quad (60)$$

Considerando-se o uso da função de ativação não-linear, as equações do gradiente descendente são modificadas para acomodar a nova função erro, eliminando-se da equação (56) o fator que contém a derivada, resultando em:

$$\delta_i^{p,k} = (d_i^{p,k} - O_i^{p,k}). \quad (61)$$

O resultado passa a ser o de incorporar pequenas correções nas regiões planas da superfície de erro e, desta forma acelerar a convergência.

C.4.1.2 – Variação da taxa de aprendizado

No método *backpropagation* padrão, a taxa de aprendizado é mantida constante durante todo o cálculo, no entanto, pode-se acelerar a convergência, adaptando a taxa de aprendizagem, de acordo com o erro total obtido, conforme a sistemática descrita a seguir.

Inicialmente define-se uma tolerância de variação para o erro e a taxa de aprendizado é feita tão grande quanto possível e passa a responder, em sua variação, de acordo com a complexidade da superfície do erro. Considerando o erro antes e após cada iteração, faz-se o acompanhamento:

- a) se a mudança no valor do peso aumentou o erro quadrado global, mais que a tolerância pré-definida, a atualização dos pesos da última iteração é descartada, a taxa de aprendizado é multiplicada por um fator que pode variar entre 0 e 1 e o momento é igualado a zero;

- b) se a mudança no valor do peso reduziu o erro quadrado global, a atualização dos pesos da última iteração é aceita e a taxa de aprendizado é multiplicada por um fator maior que 1. Se o momento, na iteração anterior foi feito igual a zero, terá seu valor restaurado;
- c) se a mudança nos pesos aumentou o erro quadrático dentro da tolerância pré-definida, então a atualização dos pesos é efetuada, mas a taxa de aprendizado e o momento permanecem com seus valores inalterados.

Se a taxa de aprendizado for feita muito alta, o algoritmo pode oscilar e tornar-se instável. Se for muito pequena, o algoritmo pode levar muito tempo para convergir.

Este método é sensível à chamada paralisia de treinamento, quando são utilizadas funções de transferência sigmoidais, nas camadas intermediárias. Derivadas muito pequenas de tais funções podem causar pequenas reduções no erro de treinamento, que são proporcionais às derivadas da função sigmoideal, fazendo com que o processo de treinamento permaneça paralisado.

C.4.1.3 – O método *resilient backpropagation*

Esta variante torna o método *backpropagation* bem mais rápido que o algoritmo que utiliza o gradiente descendente padrão, porque evita a paralisia da rede ao considerar apenas o sinal da derivada para determinar a direção da atualização dos pesos; o valor da derivada não tem efeito sobre a atualização dos pesos. Se após duas iterações sucessivas, a derivada da função em relação ao peso tiver o mesmo sinal, o valor do peso é atualizado por um delta. O

valor será reduzido também por um delta, se a derivada em relação a este peso mudar o sinal a partir da iteração anterior. Se a derivada for zero, o valor do peso continuará sendo o mesmo. Caso aconteçam oscilações nos valores dos pesos, as mudanças serão reduzidas. Se os pesos continuarem a mudar na mesma direção, após várias iterações, então, a magnitude das mudanças nos pesos será incrementada.

C.4.1.4 – O método do gradiente conjugado

Embora o algoritmo *backpropagation* básico ajuste os pesos na direção do gradiente negativo, esta é a direção na qual a função decaí mais rapidamente, mas não necessariamente, implica em uma convergência mais rápida. Na técnica do gradiente conjugado, pode-se obter grande rapidez já que este método faz pesquisa em linha usando uma série de direções conjugadas, de forma a determinar as direções que produzirão convergência mais rápida. Uma direção inicial pode ser o negativo do gradiente.

Na maior parte dos métodos baseados em gradientes conjugados, o tamanho do passo é ajustado em cada iteração. Uma desvantagem que se contrapõe ao fato deste algoritmo convergir em poucas iterações é o fato de requerer muitos cálculos por iteração, mais do que outras técnicas.

Existem variações deste algoritmo, a saber: atualizações de Fletcher-Reeves, Polak-Ribière, Powell-Beale e o gradiente conjugado escalonado.

C.4.1.5 – Métodos quasi-Newton

Esta é uma das técnicas que introduziram melhorias no método do gradiente descendente. Baseia-se no método de Newton, que encontra um mínimo para uma função não linear em um intervalo, através do mínimo encontrado para a uma função quadrática aproximadora. O ponto mínimo é encontrado onde a derivada da função quadrática aproximadora se iguala a zero. As soluções para funções não quadráticas são encontradas dentro de uma faixa de tolerância pré-definida como critério de parada e aplicando, de forma iterativa, as aproximações quadráticas locais, utilizando a série de Taylor.

Para maior entendimento, suponha que se deseja encontrar o mínimo para uma função quadrática $f(x)$. Fazendo a expansão de Taylor em torno de um ponto x_k , obtém-se:

$$f(x) = f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2}(x - x_k)f''(x_k)(x - x_k) \quad (62)$$

Achando a derivada e igualando a zero, tem-se que:

$$f'(x) = f'(x_k) + f''(x_k)(x - x_k) = 0 \quad (63)$$

Resolvendo para x :

$$x = x_k - [f''(x_k)]^{-1}f'(x_k) \quad (64)$$

Sendo a função, quadrática, o mínimo foi encontrado em x . Caso a função não seja quadrática, introduz-se o processo iterativo, fazendo $x = x_k$ e calculando um novo mínimo para x_{k+1} . Assim:

$$x_{k+1} = x_k - [f''(x_k)]^{-1}f'(x_k) \quad (65)$$

$$x_{k+1} = x_k - A_k^{-1}\nabla f(x_k) = x_k - A_k^{-1}g_k \quad (66)$$

Onde, o termo $A_k^{-1}g_k$ é um vetor que vai do ponto inicial até a solução. O ângulo é dado pela inversa da matriz de derivadas segundas, chamada Hessiana A_k^{-1} e o tamanho do segmento é dado pelo gradiente g_k .

Calcular a inversa da Hessiana envolve dificuldades e tempo computacional. Assim, surgiram os métodos quasi-Newton que fazem aproximações da Hessiana para minimizar as dificuldades de cálculo de sua inversa. Assim, os métodos Quasi-Newton não requerem o cálculo de derivadas segundas. Utilizam duas etapas: uma, de atualização da Hessiana, donde se obtém a direção de busca do mínimo e a outra etapa que, utilizando busca em linha, encontra o ponto mínimo desejado: encontrado o ponto estacionário, o sentido da concavidade é dado pela derivada segunda positiva. A Hessiana mantém-se definida positiva, indicando o ponto de mínimo.

O tamanho do passo por aproximação em cada iteração é encontrado pela busca em linha, procedimento que pode ser exemplificado, citando-se Fibonacci e o método da Seção Áurea.

Os métodos quasi-Newton convergem mais rápido que dos métodos baseados em gradientes conjugados.

Para determinar-se a atualização da Hessiana, tem-se utilizado muito o método BFGS (Broyden, 1969 e Fletcher, Goldfarb e Shanno, 1970), que garante bom desempenho.

Definindo-se:

$$s_k = x_{k+1} - x_k \quad (67)$$

e

$$\eta_k = g_{k+1} - g_k \quad (68)$$

a fórmula de atualização de BFGS é dada por:

$$\mathbf{H}_{k+1} = \mathbf{H}_k + \frac{\eta_k}{\eta_k^T \mathbf{s}_k} \frac{\eta_k \eta_k^T}{\mathbf{s}_k^T \mathbf{H}_k \mathbf{s}_k} - \frac{\mathbf{H}_k^T \mathbf{s}_k \mathbf{s}_k^T \mathbf{H}_k}{\mathbf{s}_k^T \mathbf{H}_k \mathbf{s}_k} \quad (69)$$

C.4.1.6 - Método de Levenberg-Marquardt

Este algoritmo utiliza técnicas de otimização numéricas na soma de erros quadráticos que define a função custo. O objetivo deste método é, da mesma forma que os métodos quasi-Newton, evitar o cálculo da inversa da matriz Hessiana.

As redes *feedforward* utilizam, geralmente, funções custo definidas como função da soma de quadrados, assim sendo, Levenberg-Marquardt propõem a aproximação da matriz Hessiana, através de:

$$\mathbf{H} = \mathbf{J}^T \mathbf{J} \quad (70)$$

e que o gradiente seja calculado como:

$$\mathbf{g} = \mathbf{J}^T \mathbf{e} \quad (71)$$

Onde: $\underline{\mathbf{J}}$ é a matriz Jacobiana, contendo das derivadas primeiras dos erros da rede, em relação aos pesos e limiares;

$\underline{\mathbf{e}}$ é um vetor de erros da rede.

O cálculo da matriz Hessiana é evitado pelo uso da matriz de Jacobianos, cujo cálculo pode ser feito utilizando o algoritmo *backpropagation* padrão. Este método é muito rápido e também pode incorporar recurso para adaptar o tamanho do passo.

A aproximação da matriz Hessiana pode ser utilizada, no mesmo formato das equações baseadas no método de Newton:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \left[\mathbf{J}^T \mathbf{J} + \mu \mathbf{I} \right]^{-1} \mathbf{J}^T \mathbf{e} \quad (72)$$

C.5 – Arquitetura: Memória Associativa Bidirecional (BAM)

As redes recorrentes, isto é, as que possuem laços de realimentação entre as suas saídas e suas entradas, são incondicionalmente estáveis, segundo Wasserman (1989). Além disto, a resposta de tais redes é dinâmica, sendo que a saída é recalculada e retroalimentada para modificar a entrada, logo depois de apresentada uma nova entrada. Constitui-se em memórias associativas se os pesos forem selecionados para produzir energia mínima, segundo a função de Liapunov. Assim como a memória humana é associativa, a memória da rede neural manifesta-se quando a rede é capaz de associar a estímulo de entrada, um padrão previamente armazenado.

Kosko que em 1988, desenvolveu uma memória associativa bidirecional, a BAM, também provou que ela é incondicionalmente estável, ou seja a rede não se tornará instável, qualquer que seja o conjunto de pesos. A BAM aceita um vetor de entrada sobre um conjunto de neurônios e produz uma saída relacionada, porém diferente, sobre outro conjunto de

vetores de saída. A arquitetura da BAM não é uma matriz, mas uma rede neural de duas camadas: FA e FB. É capaz de fazer generalizações, produzindo saídas corretas apesar de entradas incompletas (Wasserman, 1989). Esta é uma importante característica dessa rede. Ela realiza a associação entre pares de padrões, utilizando o princípio de aprendizado de Hebb, o biólogo que em 1949 propôs um princípio, pelo qual o aprendizado em sistemas nervosos complexos poderia ser feito considerando que as intensidades das conexões sinápticas - o peso - entre duas células seriam modificadas em acordo com o grau de atividade correlacionando tais células.

Segundo Wasserman (1989), a estrutura da BAM pode ser ilustrada através da FIG. 42. A camada 0 não tem função computacional, apenas atua na distribuição das saídas de volta às entradas. A fase de treinamento da BAM, durante a qual a rede neural faz ajustes nos pesos para aprender a associar os k pares de padrões A_k e B_k em cada direção, requerem manipulação matricial: os vetores A e B são obtidos através da apresentação, ora de um, ora de outro sinal às matrizes de pesos W e W^T , de forma repetitiva, até que a rede chegue a um ponto estável, quando não haverá mais mudanças nos vetores.

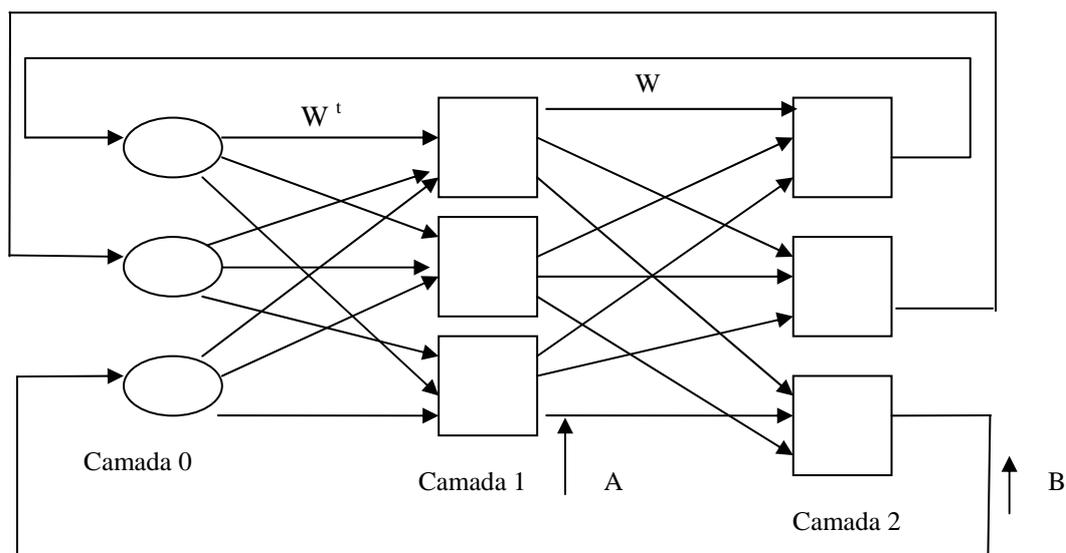


FIGURA 42 – Ilustração da estrutura da Memória associativa bidirecional – BAM, similar à rede Hopfield.
FONTE: Wasserman, 1989, p. 114

As repetidas passagens por sobre o laço que liga a entrada e a saída constitui a chamada ressonância, que só termina quando o ponto estável é encontrado e as mudanças cessam.

Diferentemente de outras redes, a matriz de pesos da BAM é construída através da soma das matrizes produtos de cada par entrada-saída. Os padrões de entrada e de saída são considerados como vetores coluna para produzir uma matriz chamada matriz mestra obtida tomando o produto do vetor A_k transposto pelo vetor B_k (produto externo dos dois vetores).

A rede tem a característica especial de fazer associações, o que permite que o chamado aprendizado seja realizado através dos valores guardados na matriz chamada mestre, que é uma matriz de pesos W , “ $n \times m$ ” definida como:

$$W = \sum_{k=1}^m A_k^T B_k \quad (73)$$

Esta matriz é ativada para responder a um padrão novo que necessite ser caracterizado, bem como para agregar novas características através da admissão de novos padrões individuais ou congregados em um arquivo.

Wasserman (1989) apresenta interessantes diversificações para a BAM: (a) a BAM não homogênea que possui limiares diferentes de zero em cada neurônio, permitindo que a rede mantenha cerca de 2^n estados estáveis; (b) BAM contínua cujos neurônios produzem suaves e contínuas respostas, de forma assíncrona, o que torna cada neurônio livre para mudar seu estado em qualquer momento, assim que suas entradas indiquem que isto deve ser feito; (c) BAM adaptativa: ajusta seus pesos durante a operação, de forma a entrar em ressonância, reduzindo sua própria energia; (d) a BAM competitiva: caracteriza-se pela implementação de interconexões entre os neurônios internos a cada camada, por meio de simétricos pesos adicionais.

C.6 – Arquitetura: Rede Hopfield

A configuração da rede de Hopfield é bastante singular, já que é composta por uma única camada com realimentação. Não há qualquer camada escondida.

Esta arquitetura foi proposta por John J. Hopfield em 1982. Seu modelo original consiste em estocar a informação através dos chamados “atratores” que são pontos fixos, definidos pelos valores dos limiares, pesos e vetores de padrões.

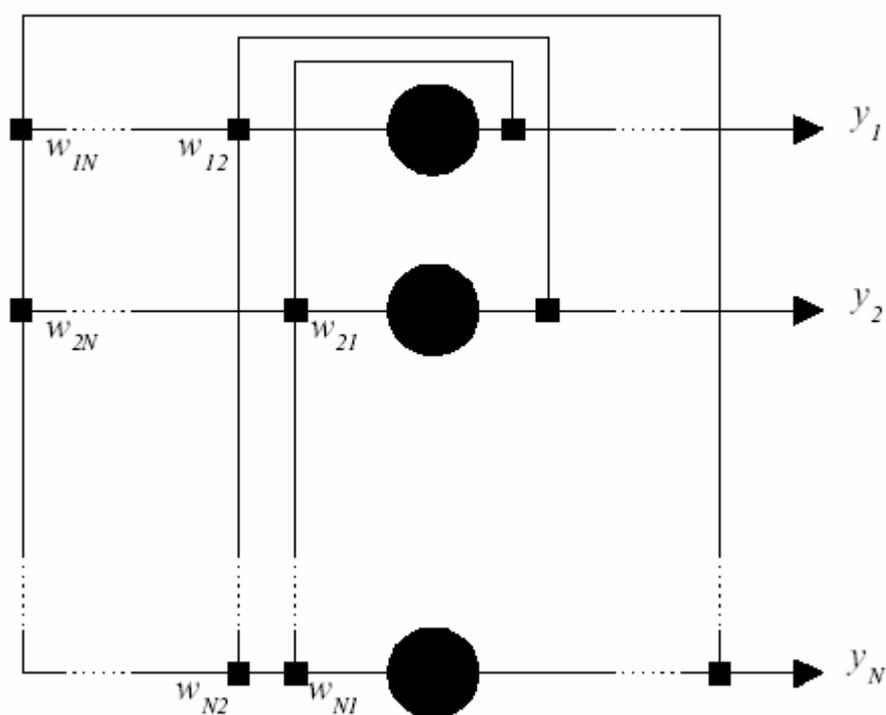


FIGURA 43 – Estrutura da Rede Neural Hopfield.
FONTE: Araújo, 1999. p. 65.

Na FIG. 43, os círculos negros representam os neurônios que tem a mesma estrutura do neurônio de McCulloch-Pitts, sendo, utilizada a função de ativação degrau, o que ocasiona saídas com valores discretos, razão pela qual a rede é conhecida por discreta ou binária. Ainda

na FIG. 43, os quadrados indicam os pesos. Hopfield propôs que a ordem de atualização dos valores das saídas dos neurônios fosse determinada randomicamente sendo, em cada instante de tempo, atualizado um único neurônio. Estando excluída a auto-realimentação, ou seja, as saídas dos neurônios não são reaplicadas às suas próprias entradas. Funcionam, portanto, de uma forma não supervisionada.

Suas principais aplicações são como memória de conteúdo endereçável e para resolver problemas de otimização combinatória - Menezes (1998). Além disto, tem demonstrado tolerância a ruídos na entrada e a dinâmica de interação dos neurônios na rede Hopfield pode ser representada por uma função de Liapunov.

A aplicação como memória associativa é viabilizada pelo fato do sistema ter a propriedade de associar um estímulo externo a um padrão previamente armazenado. Bem dimensionados, em função do problema e, funcionando perfeitamente, as chamadas “bacias de atração” vinculadas à área de influência de cada ponto, definido como atrator, promove órbitas de atração para os padrões submetidos à rede.

A rede Hopfield possui várias topologias: alguns autores como Wasserman (1989) e Kartalopoulos (1996) apresentam-na como possuindo 2 camadas: uma que recebe os sinais de entrada e as de realimentação e outra que gera o sinal de saída e de realimentação. Outros autores como Santana, Nascimento e Sakane (2003) e Araújo (1999) apresentam uma topologia diferente em que a rede possui uma única camada com “N” unidades.

Universalmente, a Hopfield é classificada como uma rede recorrente, porém, são encontradas na literatura nos modelos que possuem uma única camada:

- a) neurônios que apresentam auto-realimentação;
- b) modelos com realimentação para todas as entradas, exceto as próprias.

Esta construção levanta algumas questões, pois como consequência da própria arquitetura, uma alteração da saída de um único neurônio repercute-se nas entradas de todos

os outros. Estas alterações podem, por sua vez, provocar novas alterações nas saídas destes neurônios e assim sucessivamente. Põe-se, então, a questão de saber se a rede converge para um estado final ou se, pelo contrário, diverge indefinidamente.

Hopfield demonstrou que é possível garantir a convergência da rede se:

- a) $w_{ij} = w_{ji}$, ou seja, se as interações entre os neurônios forem simétricos;
- b) $w_{ii} = 0$, ou seja, se os neurônios não se re-alimentarem;
- c) se for utilizada a dinâmica de atualização assíncrona, ou seja, se apenas um nó for atualizado por vez.

Existem outras propostas de atualização dos valores das saídas: seqüencialmente, respeitando uma ordem previamente estabelecida entre os neurônios ou então, fazer as atualizações em paralelo. A atualização só é concluída quando todos os neurônios da rede estabilizam suas respectivas saídas. No entanto, a atualização é feita em paralelo, não garante a convergência.

Shapiro (2003) apresenta três formas de modelagem da rede Hopfield:

- a) Modelo binário: considerada de tempo discreto, a idéia-chave consiste em armazenar nos pesos das ligações entre neurônios um conjunto limitado de padrões, compostos por dígitos binários (*bits*), cujo comprimento deverá ser igual ao número de neurônios da rede. A cada dígito binário corresponderá um neurônio.
 - b) Modelo estocástico: pode ser considerado como uma variação do modelo binário, que utiliza a estatística e o cálculo de probabilidades no treinamento e processamento da rede.
 - c) Modelo contínuo: as redes de Hopfield contínuas apresentam diferenças na dinâmica dos neurônios, que passa a ser mais complexa. As saídas da rede são contínuas, o que obriga à alteração na função de ativação, que passa a ser sigmóide. Em termos de
-

aplicações práticas, as redes de Hopfield contínuas servem, normalmente, para resolver problemas de otimização. Possuem elevada velocidade.

Outras características apontadas pelos diversos autores consultados são:

- a memória associativa Hopfield opera em duas fases: uma de estocagem de dados e outra de recuperação. Daí que alguns cheguem a afirmar que não há treinamento - Araújo (1999). Outros consideram a fase de estocagem como sendo um aprendizado estático diferente das redes *Feedforward*, nas quais o aprendizado é considerado iterativo; coerentemente, no caso da Hopfield, a recuperação da informação é classificada, neste caso, como um processo dinâmico, enquanto na *Feedforward*, o processo é classificado como estático, conforme Vellasco (2003);
- em termos de função de transferência não-linear, o funcionamento da Hopfield como memória associativa requer uma função degrau, enquanto seu funcionamento em problemas de otimização, requer redes Hopfield contínuas com função de ativação sigmoidal.

Alguns aspectos devem ser tidos em conta para que uma rede Hopfield seja estável:

- 1) deve haver convergência, ou seja, a saída dos neurônios deve permanecer estável, na etapa de recuperação de informações;
 - 2) quando for submetido um padrão de entrada incompleto ou incorreto, a rede deve estabilizar em um “atrator” mais próximo ao padrão apresentado. O padrão correspondente ao “atrator”, passa a ser considerado como padrão correto, a resposta da rede. Pode ocorrer que a rede ao receber padrões corrompidos, não apresente convergência para um estado final que corresponda ao padrão que mais se aproxima do padrão aplicado inicialmente. É importante referir que isto nem sempre acontece e que a memória pode evocar um padrão inexistente (padrões
-

espúrios), ou outro padrão, que não o mais semelhante (perda de memória), além disto, nem todas as memórias são evocadas com igual ênfase.

C.6.1 - Considerações de Desempenho

O desempenho da rede depende do número máximo de padrões armazenável numa rede com neurônios. As Redes de Hopfield apresentam bom desempenho como memória associativa imune a ruído quando: (a) o número de neurônios da rede é grande; (b) os padrões de entrada são ortogonais.

Quanto à sensibilidade ao ruído: se a diferença entre o número de neurônios da rede e o número de padrões é muito grande, o ruído fica desprezível, senão, o ruído torna-se mais significativo.

Várias bibliografias, dentre elas: Jain, Mao e Mohuddin (1996) e Araújo (1999) - referindo Dayhoff (1996) - afirmam que: a capacidade de memorizar da Hopfield é finita e que o número máximo de padrões randômicos que uma rede Hopfield pode estocar é $0,15 n$, onde “n” é o número de unidades binárias (neurônios).

Outros afirmam que esta capacidade de $0,15 n$ é pessimista e que para haver sobrecarga de padrões, é necessário que a relação entre o número de neurônios da rede e o número de padrões seja da ordem de 50%, quando a rede passa a apresentar estados espúrios de equilíbrio e a ter maior possibilidade de convergir para padrões diferentes do desejado.

Wasserman (1987) cita Cohen e Grossberg (1983), em relação à demonstração de que uma rede recorrente é sempre estável, se a matriz é simétrica, com zeros sobre a diagonal principal.

Araújo (1999) demonstra que: (a) a determinação dos pesos das ligações é efetuada com base no princípio Hebbiano; (b) que as redes de Hopfield binárias convergem para um estado mínimo e estável de energia, representando-se a dinâmica de interação dos neurônios por uma função de Liapunov.

As redes BAM e Hopfield possuem uma série de características em comum: (a) ambas são classificadas como recorrentes, por possuírem conexões ponderadas (pesos) em realimentação entre a saída e as camadas anteriores. (b) Apresentam limitações quanto à quantidade de padrões armazenados em relação ao número de neurônios. (c) Possuem duas fases: uma de estocagem e outra de recuperação de dados. (d) Ambas são consideradas memórias associativas. (e) Da mesma forma que a *feedforward*, utilizam a regra de Hebb. (f) Possuem processo de estocagem de dados considerado como um aprendizado não supervisionado. (g) Segundo Wasserman (1989), caso a matriz de pesos seja feita quadrada e simétrica e o número de camadas seja reduzido a uma, a BAM torna-se a uma rede Hopfield auto-associativa.

C.7 – Arquitetura: Rede de Base Radial (RBF)

A rede RBF é uma classe especial de rede *feedforward* multicamada, que foi desenvolvida por M.J.D. Powell em 1985. É capaz de fazer classificação e de funcionar como aproximador universal de funções, desde que haja um número suficiente de neurônios na camada escondida.

A forma básica da rede função de base radial de acordo com Haykin (2001) envolve três camadas, a saber: a camada de entrada, constituída por nós que conectam a rede ao seu ambiente; uma única camada oculta e a camada de saída que é linear.

Ainda segundo Haykin (2001), que cita Powell (1988), a técnica consiste em escolher uma função da forma

$$F(\mathbf{x}) = \sum_{i=1}^N w_i \varphi(\|\mathbf{x} - \mathbf{x}_i\|) \quad (74)$$

onde $\{ \varphi_{ij} = \varphi(\|\mathbf{x}_j - \mathbf{x}_i\|), (j,i) = 1,2,\dots,N \}$ é um conjunto de N funções arbitrárias e $\| \cdot \|$

representa uma normal euclidiana. Os pontos de dados conhecidos $\mathbf{x}_i \in \mathfrak{R}^{m_0}$, $i = 1,2,\dots,N$ são tomadas como centros das funções de base radial, considerando ainda os vetores:

“d” → vetor resposta desejada

“w” → vetor de peso linear

obs: ambos os vetores são de dimensão $N \times 1$ (tamanho da amostra de treinamento) e que:

$$\Phi = \{ \varphi_{ij} = \varphi(\|\mathbf{x}_j - \mathbf{x}_i\|), (j,i) = 1,2,\dots,N \} \quad (75)$$

é a matriz de interpolação.

A equação pode ser apresentada na seguinte forma:

$$\Phi \cdot \mathbf{w} = \mathbf{x} \quad (76)$$

onde: Φ é não singular, ou seja, possui inversa Φ^{-1}

e o vetor de pesos \mathbf{w} pode ser obtido através de:

$$\mathbf{w} = \Phi^{-1} \cdot \mathbf{x} \quad (77)$$

A condição de não singularidade adequada ao uso em redes RBF só pode ser obtida se os pontos $\{x_i\}_{i=1}^N$ forem todos diferentes.

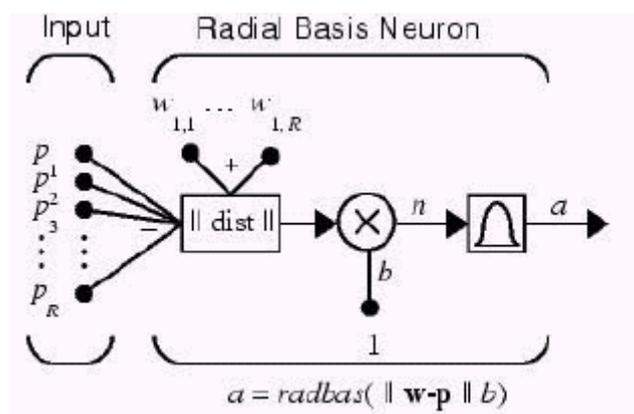


FIGURA 44 – Arquitetura da Rede Neural de Base Radial – RBF.
Ref.: Help do Toolbox de Redes Neurais do MATLAB versão 6.5.

A FIG. 44 apresenta uma rede neural de base radial com um modelo de neurônio que é diferente do apresentado até aqui (FIG. 39). O funcionamento do elemento processador também é diferente: é calculada a distância entre o vetor de entrada e o vetor de pesos, bem como o termo *bias* (“b”) que ajusta a sensibilidade do neurônio. O produto ponto-a-ponto entre o vetor diferença e o termo *bias* (limiar) é submetido à função de transferência de base radial, na camada escondida. A idéia é que os vetores de entrada e pesos definam pontos no espaço N-dimensional, de forma que a resposta do neurônio diminua rapidamente conforme esses pontos se distanciem do centro de cada função radial. O conjunto de neurônios

escondidos é projetado de forma que as suas respostas cubram todas as regiões significativas do espaço de vetores de entrada.

O funcionamento da RBF, em classificação de padrões, é feito, tomando o conjunto de vetores pesos associados com os nós, cujas funções de base radial formatam a sua resposta individualmente. Tanto as posições centrais quanto as larguras laterais da parte central dessas funções são aprendidas a partir dos padrões de treinamento. Cada unidade de saída implementa uma combinação destas funções de ativação radiais.

As funções radiais das camadas ocultas atuam como detectores que produzem saídas com valor “1” se o vetor de entradas “p” é idêntico ao vetor de pesos “w”. O reconhecimento é feito, tomando como base o grau de similaridade na forma dos padrões e não em posicionamento ou tamanho dos padrões.

Para melhor entendimento do funcionamento de uma rede neural de base radial – RBF- apresentaremos a função da forma gaussiana, utilizada neste tipo de rede como função de ativação nos elementos de processamento da rede. A FIG. 45 apresenta a função gaussiana correspondente à equação (78), com suas principais características: ser simétrica em relação à origem, o que requer um valor de desvio padrão pré-definido, aqui representado por “u”, maior que zero, para dar forma à função gaussiana e onde o valor de “x” no ponto central, corresponde ao valor máximo da função $h(x)$. A resposta para outros valores da abscissa decai rapidamente conforme os valores afastem-se do ponto central. O chamado “campo receptivo” define a faixa significativa dos valores de “x”.

$$h(x) = \exp\left(-\frac{x^2}{u}\right) \quad (78)$$

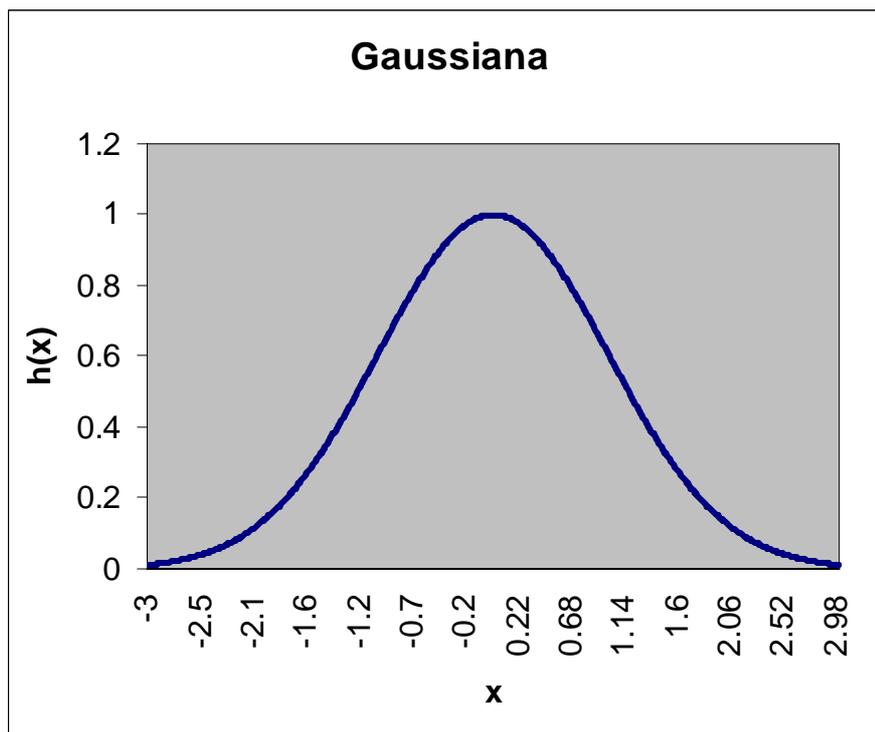


FIGURA 45 – Exemplo de uma função Gaussiana gerada a partir da equação (78).

C.7.1 – Topologia característica

A FIG.46 apresenta as duas camadas de processadores da rede característica RBF: a camada intermediária, que recebe o vetor de entrada e que utiliza funções de ativação de base radial e a segunda camada que é a de saída e utiliza função de ativação linear.

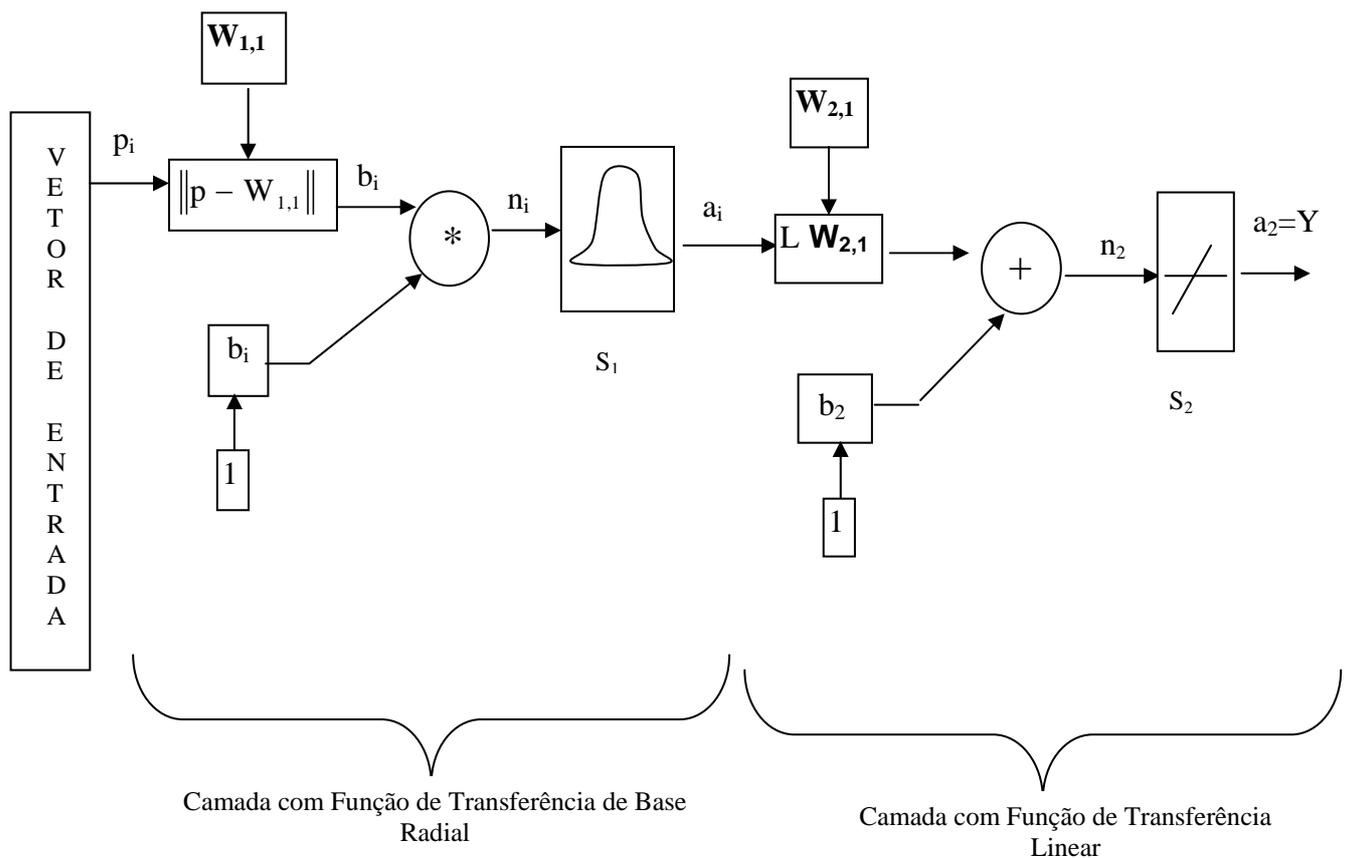


FIGURA 46 – Arquitetura da Rede Neural de Base Radial – RBF.
 Ref.: Help do Toolbox de Redes Neurais do MATLAB versão 6.5.

Para melhor entendimento da FIG. 46, considere-se:

- b_1 = vetor com os limiares pertencentes aos nós camada Radial – função S_1 ;
- b_2 = vetor com os limiares pertencentes aos nós camada Linear – função S_2 ;
- “ $W_{1,1}$ ” = vetor de pesos da camada de base radial;
- “ $W_{2,1}$ ” = vetor de pesos da camada linear;
- “ X_i ” = vetor de entrada com p_i padrões de entrada/saída;
- a_1 = posição do vetor de saída da camada de base radial;
- a_2 = posição do vetor de saída da camada linear;
- S_1 = função de transferência de base radial;
- S_2 = função de transferência linear;

- n_1 = posição do vetor de entrada na função de transferência de base radial;
- n_2 = posição do vetor de entrada na função de transferência linear.
- Y = vetor de saída da rede RBF.

A camada oculta é composta por um conjunto de neurônios cujas funções de transferência de base radial recebem o vetor distância entre o vetor de pesos " $\mathbf{W}_{1,1}$ " e o vetor de entrada " \mathbf{X}_1 ", que é multiplicado elemento-a-elemento pelo vetor de limiares " \mathbf{b}_1 ", gerando o produto " \mathbf{n}_1 " que é, então, submetido às funções gaussianas. Por sua vez, a saída " \mathbf{a}_1 ", multiplicada pelo vetor de pesos da camada de saída " $\mathbf{W}_{2,1}$ " e somada ao vetor de limiares " \mathbf{b}_2 ", compõe a entrada a ser submetida à equação linear da camada de saída, gerando a saída " \mathbf{Y} ".

O comportamento da rede RBF é muito interessante: (a) se os neurônios da camada de base radial tiverem vetores de pesos muito diferentes do vetor de entradas, o vetor de saída, após a função radial, terá saídas próximas de zero, sendo tais saídas efeitos desprezíveis sobre os neurônios da camada linear; (b) um neurônio de base radial que possua um vetor de pesos bem próximo do vetor de entradas, terá um vetor de saída após a função de transferência radial, com valor próximo de 1, valores estes que serão submetidos através da camada linear; (c) os neurônios possuem pesos que, normalmente, possuem distâncias variadas em relação aos padrões de entrada, o que causará disparos variando em graus, em relação à camada linear.

Cada neurônio da camada intermediária calcula uma função de base radial onde:

- a) o centro - μ - da função radial é definido em função do vetor de pesos;
- b) a largura da função de base radial - σ - define sua área de influência, seu campo receptivo.

Corresponde ao desvio médio padrão.

A função de ativação da camada escondida é não-linear, cujo valor aumenta ou diminui com a distância ao ponto central da função radial.

Uma função típica utilizada é:

$$\phi(v) = \exp\left(-\frac{v^2}{2\sigma^2}\right) \quad (79)$$

Onde:

$$v = \|x - \mu\| \quad (80)$$

x = vetor com os padrões de entrada

μ = centro da função radial

σ = Largura da função de base radial

C.7.2 – Treinamento da rede RBF

Existe uma variedade de algoritmos de aprendizagem para redes RBF's. As técnicas básicas empregam uma estratégia híbrida de aprendizado em 2 passos. O primeiro estima o número de neurônios da camada intermediária, a partir das informações sobre o conjunto de dados; seleciona os parâmetros das funções gaussianas, utilizando um algoritmo não supervisionado para determinar os pesos entre as camadas ocultas. Para determinar os pesos entre as camadas oculta e de saída utiliza o algoritmo supervisionado baseado no erro médio quadrático (LMS), permissível pelo fato das unidades de saída serem lineares.

Esse algoritmo híbrido converge mais rápido que o algoritmo *backpropagation* padrão para treinamento de Perceptrons multicamadas, porém, o fato da rede RBF possuir, frequentemente, grande quantidade de nós na camada oculta, torna-a de processamento mais lento que a *backpropagation*.

A RBF define hiper-elipsóides como forma de subdividir o espaço de entrada, a partir das características da camada intermediária.

Durante o treinamento, são estabelecidos os parâmetros ajustáveis da rede. Devem ser especificados valores para: (1) os centros das funções radiais (μ_i); (2) os desvios padrões (σ_i) para cada neurônio da camada escondida, além dos vetores peso (W_i), de forma a minimizar o erro médio entre a saída da rede e a saída desejada.

A determinação dos centros das funções radiais (μ_i) é uma questão extremamente importante para o bom funcionamento da rede e existem várias alternativas para a sua determinação.

- a) Método mais simples: definir um centro para cada vetor de entrada do conjunto de treinamento, o que cria mais neurônios do que o realmente necessário (*overfitting*);
 - b) Método mais complexo: determina-se o centro de cada grupo de vetores de entrada, estabelecendo um processador para cada conjunto de padrões de entrada com distâncias equivalentes em relação a cada centro de agrupamento (vetores pesos);
 - c) Outros dois métodos: fazer seleção aleatória a partir dos valores de treinamento, considerando-os representativos do problema e definir os centros, considerando-os dispostos em uma grade com distâncias uniformes entre dois centros vizinhos.
-

A determinação dos desvios padrões (σ_i) que definem o diâmetro do Campo Receptivo de cada neurônio pode ter um grande efeito na precisão do sistema. Vários métodos podem ser usados:

- a) considerar que todas as larguras são iguais à média de todas as distâncias Euclidianas entre os centros de unidades vizinhas.
- b) atribuir a cada unidade, uma largura obtida a partir da distância entre o seu centro e o centro da unidade mais próxima.
- c) atribuir a cada (σ_i), a distância média de seu centro aos N vetores de entrada mais próximos.
- d) atribuir a cada (σ_i) um valor constante e igual a 1.

O treinamento dos pesos da camada de saída é supervisionado, buscando, no processo, que os valores de saída da rede aproximem-se dos valores esperados, conforme o erro máximo permitido. Cada vetor de entrada do conjunto de treinamento é aplicado à entrada da rede, gerando o cálculo das saídas dos processadores da camada escondida. As saídas são comparadas aos valores de saída esperados, gerando um ajuste nos pesos que objetiva reduzir tais diferenças. A cada iteração, novos ajustes nos pesos são feitos até que o erro seja pequeno.

Destacaremos duas variações das redes de bases radiais, a saber: a GRNN - *Generalized Regression Neural Network* – e a rede neural probabilística – PNN.

C.7.3 – A Rede Neural de Regressão Generalizada - GRNN

Baseia-se na Teoria de Regressão Não-Linear: calcula-se o valor mais provável de y , para dado x e um certo conjunto de treinamento, que pode conter ruído, considerando-se a regressão de uma variável dependente y sobre uma variável independente x .

Esta rede possui algumas características que a distinguem: (a) sua topologia possui múltiplas saídas; (b) não requer treinamento iterativo; (c) aproxima qualquer função entre entrada e saída (a partir dos dados de treinamento), o que possibilita que funcione também como uma rede classificadora; (d) conforme o conjunto de treinamento cresce, o erro se aproxima de zero.

A FIG. 47 destaca a função de normalização que caracteriza a GRNN: a saída de cada processador da camada de saída é dividida pela soma de todas as saídas dos processadores da camada escondida.

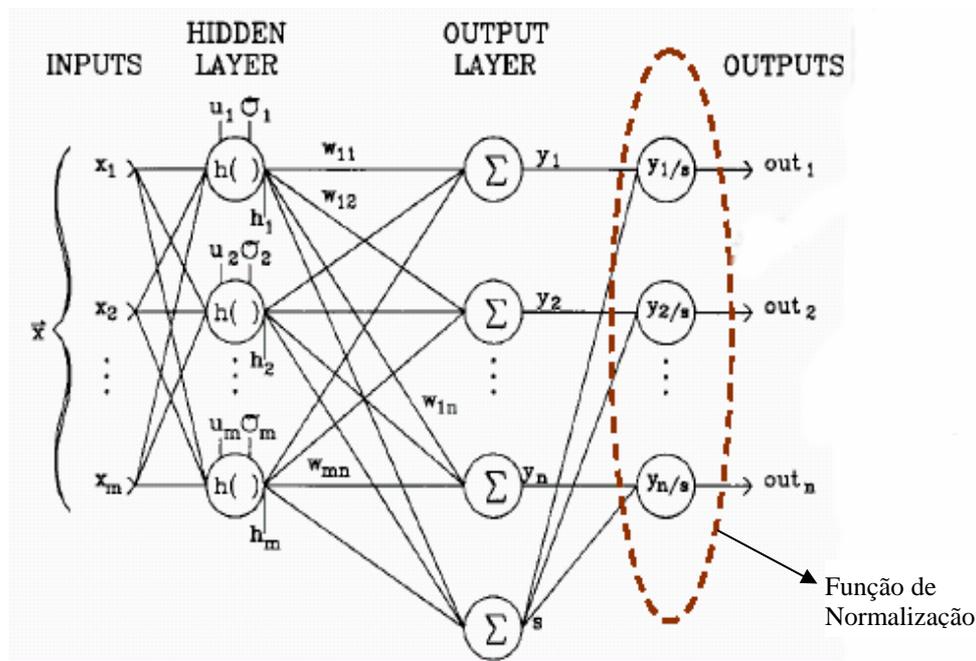


FIGURA 47 – Arquitetura da Rede Neural de Base Radial tipo GRNN.
Ref.: Help do Toolbox de Redes Neurais do MATLAB versão 6.5.

C.7.4 - Topologia das Redes Neurais PNN - Probabilistic Neural Network Architecture

Redes Neurais Probabilísticas são redes de duas camadas: o vetor de entrada com os padrões de entrada é recebido pela primeira camada que é a camada intermediária e que utiliza funções de ativação de base radial; a segunda camada é a de saída e utiliza função de ativação linear.

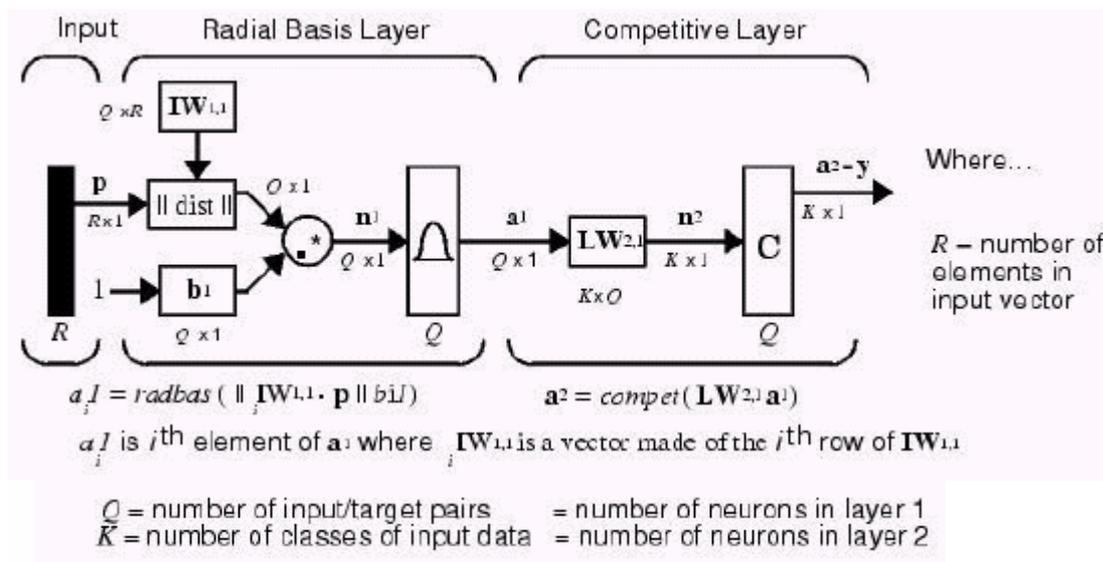


FIGURA 48 – Arquitetura da Rede Neural de Base Radial – RBF, tipo Rede Neural Probabilística. Ref.: Help do Toolbox de Redes Neurais do Matlab versão 6.5.

A rede PNN deriva dos Classificadores Bayesianos e pode, por isto, ser utilizada como uma rede Classificadora.

Quando um padrão de entrada é apresentado à rede, a primeira camada calcula as distâncias entre o vetor de entrada e os vetores do conjunto treinado. A segunda camada soma as contribuições para cada classe de entrada, produzindo, como saída um vetor de probabilidades. A função de transferência competitiva toma o máximo das probabilidades e produz uma saída em “1” para a classe de maior probabilidade, zerando as outras classes.

Podemos compara as redes GRNN e PNN: adicionando-se a camada de decisão da PNN, as redes ficam equivalentes. Além disto, o método de Regressão, no qual baseia-se a GRNN é uma estimativa ótima no sentido de se encontrar o Erro Médio Quadrático, o que também pode fazer qualquer estimador, inclusive o Classificador Bayesiano desde que com um número suficiente de dados de treinamento.

Em termos de características globais, podemos concluir que as Redes RBFs são métodos poderosos, com grande aplicabilidade, treinamento rápido e aprendizado incremental.

A comparação com as redes Perceptron Multicamadas treinadas com algoritmo *Backpropagation*, revela: (a) Maior rapidez de aprendizado para a RBF; (b) vantagem de não apresentar problemas de Paralisia da Rede e Mínimo Local; (c) fase de recuperação de dados mais lenta em relação à *Backpropagation*. (d) a RBF requer mais neurônios do que a *Backpropagation*, já que existe um inter-relacionamento entre o número de vetores de treinamento, o número de nós da camada intermediária e o melhor desvio padrão para a situação. O número de vetores de treinamento deve cobrir o espaço de entrada com densidade suficiente para alcançar a precisão desejada. O desvio padrão deve ser grande o suficiente para produzir a generalização desejada. Já a *Backpropagation* é compacta e requer menor número de neurônios, por distribuir internamente a diferenças que equilibram a rede. (e) as RBF's são mais rápidas para serem projetadas.

.
